

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2005, 00:22:11 ; Search time 1.43846 Seconds  
(without alignments)  
735.775 Million cell updates/sec

Title: US-10-049-822A-12  
Perfect score: 58  
Sequence: 1 YGRKKRQRRR 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79:.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	58	100.0	71	2	TN9384
2	58	100.0	72	1	TNLJH4
3	58	100.0	86	1	TNLJZR
4	58	100.0	86	2	JCS591
5	58	100.0	86	2	A25700
6	58	100.0	86	2	S54381
7	58	100.0	86	2	S33982
8	58	100.0	87	2	T01665
9	58	100.0	95	1	TNLJ12
10	58	100.0	101	1	E44001
11	58	100.0	101	2	T09446
12	55	94.8	86	1	TNLJND
13	46	79.3	100	1	TNLJSI
14	43	74.1	57	2	A34356
15	42	72.4	250	2	D38095
16	42	72.4	268	2	C38095
17	42	72.4	269	2	A38095
18	42	72.4	269	2	B38095
19	42	72.4	272	2	A38900
20	42	72.4	279	2	JH0402
21	42	72.4	303	2	JH0401
22	42	72.4	366	2	S61796
23	42	72.4	525	2	T48824
24	42	72.4	747	1	A57107
25	42	72.4	1193	2	A86193
26	41	70.7	37	2	S29829
27	41	70.7	47	2	F58208
28	41	70.7	47	2	E58208
29	41	70.7	58	2	S34045

30	41	70.7	294	2	E87538
31	41	70.7	367	2	S59329
32	41	70.7	417	2	E30341
33	41	70.7	517	2	A45121
34	41	70.7	770	2	G88445
35	40	69.0	78	2	A40973
36	40	69.0	79	2	S56116
37	40	69.0	91	2	A59493
38	40	69.0	118	2	S56117
39	40	69.0	165	2	A59492
40	40	69.0	371	2	B39625
41	40	69.0	397	2	A39565
42	40	69.0	399	2	A39625
43	40	69.0	952	2	E86147
44	39	67.2	46	2	A18865
45	39	67.2	90	2	B32986

ALIGNMENTS

RESULT 1

T09384  
trans-activating transcription regulator - human immunodeficiency virus type 1 (isolate C;Species: human immunodeficiency virus type 1, HIV-1  
C;Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004  
C;Accession: T09384  
R;Michael, N.L.; Chang, G.; d'ARCY, L.A.; Ehrenberg, P.K.; Mariani, R.; Busch, M.P.; Bir J. Virol. 69, 4228-4236, 1995  
A;Title: Defective accessory genes in a human immunodeficiency virus type 1-infected lon A;Reference number: Z16654; MUID:95287475; PMID:7769682  
A;Accession: T09384  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-71 <MIC>  
A;Cross-references: UNIPROT:Q71926; EMBL:U24451; NID:G829440; PIDN:AAA79576.1; PID:G8294 C;Genetics:  
A;Gene: tat  
C;Superfamily: AIDS trans-activating transcription regulator  
C;Keywords: transcription

Query Match 100.0%; Score 58; DB 2; Length 71;  
Best Local Similarity 100.0%; Pred. No. 0.024;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 YGRKKRQRRR 11  
|||||||  
Db 47 YGRKKRQRRR 57

RESULT 2

TNLJH4  
trans-activating transcription regulator - human immunodeficiency virus type 1 (isolate C;Species: human immunodeficiency virus type 1, HIV-1  
C;Date: 30-Jun-1988 #sequence\_revision 30-Jun-1988 #text\_change 02-Jul-1998  
C;Accession: B25523  
R;Desai, S.M.; Kalyanaraman, V.S.; Casey, J.M.; Srinivasan, A.; Andersen, P.R.; Devare, Proc. Natl. Acad. Sci. U.S.A. 83, 8380-8384, 1986  
A;Title: Molecular cloning and primary nucleotide sequence analysis of a distinct human A;Reference number: A94136; MUID:87041461; PMID:3490666  
A;Accession: B25523  
A;Molecule type: DNA  
A;Residues: 1-72 <DES>  
A;Cross-references: GB:M13137; NID:G326460  
A;Note: the GenBank entry ADBE3AA PID:G209908 differs from the published sequence in tra C;Genetics:  
A;Gene: tat  
C;Superfamily: AIDS trans-activating transcription regulator  
C;Keywords: transcription regulation

Query Match 100.0%; Score 58; DB 1; Length 72;  
Best Local Similarity 100.0%; Pred. No. 0.024;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YGRKKRQRRR 11  
|||||  
Db 47 YGRKKRQRRR 57

## RESULT 3

TNLJZR

trans-activating transcription regulator - human immunodeficiency virus Zr-6  
C;Species: human immunodeficiency virus Zr-6  
C;Date: 30-Sep-1987 #sequence\_revision 30-Sep-1987 #text\_change 09-Jul-2004  
C;Accession: C26192  
R;Srinivasan, A.; Anand, R.; York, D.; Ranganathan, P.; Feorino, P.; Schochetman, G.; Cu  
Gene 52, 71-82, 1987  
A;Title: Molecular characterization of human immunodeficiency virus from Zaire: nucleoti  
A;Reference number: A26192; MUID:87248097; PMID:3036660  
A;Accession: C26192  
A;Molecule type: DNA  
A;Residues: 1-86 <SRI>  
C;Cross-references: UNIPROT:P04609; GB:K03458; GB:M16322; NID:g329398; PIDN:AAA45377.1;  
C;Genetics:  
A;Gene: tat  
A;Introns: 72/3  
C;Superfamily: AIDS trans-activating transcription regulator  
C;Keywords: AIDS; immunodeficiency; transcription regulation

Query Match 100.0%; Score 58; DB 1; Length 86;  
Best Local Similarity 100.0%; Pred. No. 0.028;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YGRKKRQRRR 11  
|||||  
Db 47 YGRKKRQRRR 57

## RESULT 4

JC5591

transactivator protein - human immunodeficiency virus type 1  
N;Alternate names: tat protein  
C;Species: human immunodeficiency virus type 1, HIV-1  
C;Date: 23-Sep-1997 #sequence\_revision 23-Sep-1997 #text\_change 17-Mar-1999  
C;Accession: JC5591  
R;Hoffmann, S.; Willbold, D.  
Biochem. Biophys. Res. Commun. 235, 806-811, 1997  
A;Title: A selection system to study protein-RNA interactions: Functional display of HIV  
A;Reference number: JC5591; MUID:97350867; PMID:9207243  
A;Accession: JC5591  
A;Molecule type: protein  
A;Residues: 1-86 <HO2>  
C;Comment: This protein is a key regulatory protein in the viral replication cycle and h  
C;Superfamily: AIDS trans-activating transcription regulator  
F;22-31/Region: cysteine-rich

Query Match 100.0%; Score 58; DB 2; Length 86;  
Best Local Similarity 100.0%; Pred. No. 0.028;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YGRKKRQRRR 11  
|||||  
Db 47 YGRKKRQRRR 57

## RESULT 5

A25700

trans-activating transcription regulator - human immunodeficiency virus type 1 (isolate  
C;Species: human immunodeficiency virus type 1, HIV-1  
C;Date: 28-Sep-1987 #sequence\_revision 28-Sep-1987 #text\_change 09-Jul-2004  
C;Accession: A25700  
R;Sodroski, J.; Patarca, R.; Rosen, C.; Wong-Staal, F.; Haseltine, W.  
Science 229, 74-77, 1985  
A;Reference number: A25700; MUID:85244627; PMID:2990041  
A;Accession: A25700  
A;Status: preliminary

A;Molecule type: mRNA  
A;Residues: 1-86 <SOD>  
A;Cross-references: UNIPROT:P04610  
C;Superfamily: AIDS trans-activating transcription regulator

Query Match 100.0%; Score 58; DB 2; Length 86;  
Best Local Similarity 100.0%; Pred. No. 0.028;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YGRKKRQRRR 11  
|||||  
Db 47 YGRKKRQRRR 57

## RESULT 6

S54381

tat protein - human immunodeficiency virus type 1  
C;Species: human immunodeficiency virus type 1, HIV-1  
C;Date: 15-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 09-Jul-2004  
C;Accession: S54381  
R;Theodore, T.; Buckler-White, A.J.  
submitted to the EMBL Data Library, July 1989  
A;Reference number: S54377  
A;Accession: S54381  
A;Status: preliminary  
A;Molecule type: genomic RNA  
A;Residues: 1-86 <THE>  
A;Cross-references: UNIPROT:P12506; EMBL:M22639; NID:g329377; PIDN:AAA45363.1; PID:g3293  
C;Genetics:  
A;Introns: 72/2  
C;Superfamily: AIDS trans-activating transcription regulator

Query Match 100.0%; Score 58; DB 2; Length 86;  
Best Local Similarity 100.0%; Pred. No. 0.028;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YGRKKRQRRR 11  
|||||  
Db 47 YGRKKRQRRR 57

## RESULT 7

S33982

trans-activating transcription regulator - human immunodeficiency virus type 1  
C;Species: human immunodeficiency virus type 1, HIV-1  
C;Date: 06-Oct-1994 #sequence\_revision 23-Feb-1996 #text\_change 09-Jul-2004  
C;Accession: S33982; S19864  
R;Carlini, F.

submitted to the EMBL Data Library, November 1991

A;Reference number: S33979

A;Accession: S33982

A;Molecule type: DNA

A;Residues: 1-86 <CAR>

A;Cross-references: UNIPROT:P04606; EMBL:Z11530; NID:g60192; PIDN:CAA77625.1; PID:g60196

R;Siderovski, D.P.; Matsuyama, T.; Frigerio, E.; Chui, S.; Min, X.; Erfle, H.; Summer-Sm

Nucleic Acids Res. 20, 5311-5320, 1992

A;Title: Random mutagenesis of the human immunodeficiency virus type-1 trans-activator o

A;Reference number: S26385; MUID:93065196; PMID:1437550

A;Accession: S26385

A;Molecule type: nucleic acid

A;Residues: 1-86 <SID>

A;Cross-references: EMBL:X64650; NID:g60144; PIDN:CAA45921.1; PID:g60145

C;Genetics:

A;Gene: tat

A;Introns: 72/2

C;Superfamily: AIDS trans-activating transcription regulator

C;Keywords: AIDS; immunodeficiency

Query Match 100.0%; Score 58; DB 2; Length 86;  
Best Local Similarity 100.0%; Pred. No. 0.028;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YGRKKRQRRR 11

Db 1 YGRKKRQRRR 57  
|||||  
RESULT 8  
T01665  
tat protein - human immunodeficiency virus type 1  
C;Species: human immunodeficiency virus type 1, HIV-1  
C;Date: 19-Feb-1999 #sequence\_revision 19-Feb-1999 #text\_change 09-Jul-2004  
C;Accession: T01665  
R;Alizon, M.; Wain-Hobson, S.; Gluckman, J.C.; Sonigo, P.  
Cell 46, 63-74, 1986  
A;Title: Genetic variability of the AIDS virus: Nucleotide sequence analysis of two isolates  
A;Reference number: Z14389; MUID:86245056; PMID:2424612  
A;Accession: T01665  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-87 <ALI>  
A;Cross-references: UNIPROT:P04613; EMBL:K03456; NID:G60228; PIDN:CAA28015.1; PID:G60233  
C;Genetics:  
A;Introns: 72/2  
C;Superfamily: AIDS trans-activating transcription regulator  
Query Match 100.0%; Score 58; DB 2; Length 87;  
Best Local Similarity 100.0%; Pred. No. 0.028;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 YGRKKRQRRR 11  
|||||  
Db 47 YGRKKRQRRR 57  
RESULT 9  
TNLJ12  
trans-activating transcription regulator - human immunodeficiency virus type 1 (isolate)  
C;Species: human immunodeficiency virus type 1, HIV-1  
A;Note: host Homo sapiens (man)  
C;Date: 04-Dec-1986 #sequence\_revision 04-Dec-1986 #text\_change 09-Jul-2004  
C;Accession: A04017  
R;Arya, S.K.; Gallo, R.C.  
Proc. Natl. Acad. Sci. U.S.A. 83, 2209-2213, 1986  
A;Title: Three novel genes of human T-lymphotropic virus type III: immune reactivity of  
A;Reference number: A94093; MUID:86177573; PMID:3008154  
A;Accession: A04017  
A;Molecule type: DNA  
A;Residues: 1-95 <ARY>  
A;Cross-references: UNIPROT:P04326  
C;Genetics:  
A;Gene: tat  
C;Superfamily: AIDS trans-activating transcription regulator  
Query Match 100.0%; Score 58; DB 1; Length 95;  
Best Local Similarity 100.0%; Pred. No. 0.03;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 YGRKKRQRRR 11  
|||||  
Db 56 YGRKKRQRRR 66  
RESULT 10  
E44001  
trans-activating transcription regulator - human immunodeficiency virus type 1 (strain Y  
N;Alternate names: tat protein  
C;Species: human immunodeficiency virus type 1, HIV-1  
A;Note: host Homo sapiens (man)  
C;Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Jul-2004  
C;Accession: E44001  
R;Li, Y.; Hui, H.; Burgess, C.J.; Price, R.W.; Sharp, P.M.; Hahn, B.H.; Shaw, G.M.  
J. Virol. 66, 6587-6600, 1992  
A;Title: Complete nucleotide sequence, genome organization, and biological properties of  
A;Reference number: A44001; MUID:93021387; PMID:1404605

A;Accession: E44001  
A;Molecule type: DNA  
A;Residues: 1-101 <LIY>  
A;Cross-references: UNIPROT:P35965; GB:M93258  
C;Genetics:  
A;Gene: tat  
A;Introns: 72/2  
C;Superfamily: AIDS trans-activating transcription regulator  
C;Keywords: AIDS; immunodeficiency; transcription regulation  
Query Match 100.0%; Score 58; DB 1; Length 101;  
Best Local Similarity 100.0%; Pred. No. 0.032;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 YGRKKRQRRR 11  
|||||  
Db 47 YGRKKRQRRR 57  
RESULT 11  
T09446  
tat protein - human immunodeficiency virus type 1 (strain JRFL)  
C;Species: human immunodeficiency virus type 1, HIV-1  
C;Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 09-Jul-2004  
C;Accession: T09446  
R;Pang, S.; Vinters, H.V.; Akashi, T.; O'Brien, W.A.; Chen, I.S.; Koyanagi, Y.; Namazie,  
submitted to the EMBL Data Library, July 1996  
A;Reference number: Z16673  
A;Accession: T09446  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-101 <PAN>  
A;Cross-references: UNIPROT:Q75758; EMBL:U63632; NID:G1465777; PID:G1465783  
C;Genetics:  
A;Gene: tat  
A;Introns: 72/2  
C;Superfamily: AIDS trans-activating transcription regulator  
Query Match 100.0%; Score 58; DB 2; Length 101;  
Best Local Similarity 100.0%; Pred. No. 0.032;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 YGRKKRQRRR 11  
|||||  
Db 47 YGRKKRQRRR 57  
RESULT 12  
TNLJND  
trans-activating transcription regulator - human immunodeficiency virus type 1 (isolate)  
C;Species: human immunodeficiency virus type 1, HIV-1  
A;Note: host Homo sapiens (man)  
C;Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 09-Jul-2004  
C;Accession: JQ0071  
R;Spire, B.; Sire, J.; Zachar, V.; Rey, F.; Barre-Sinoussi, F.; Galibert, F.; Hampe, A.;  
Gene 81, 275-284, 1989  
A;Title: Nucleotide sequence of HIV1-NDK: a highly cytopathic strain of the human immuno-  
A;Reference number: JQ0065; MUID:90034200; PMID:2806917  
A;Accession: JQ0071  
A;Molecule type: DNA  
A;Residues: 1-86 <SPI>  
A;Cross-references: UNIPROT:P18804; GB:M27323; NID:G328154; PIDN:AAA44866.1; PID:G328155  
C;Genetics:  
A;Gene: tat  
C;Superfamily: AIDS trans-activating transcription regulator  
C;Keywords: AIDS; immunodeficiency; transcription  
Query Match 94.8%; Score 55; DB 1; Length 86;  
Best Local Similarity 90.9%; Pred. No. 0.079;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 YGRKKRQRRR 11  
|||||

Db 47 YGRKKRQRRK 57

## RESULT 13

TNLUSI

trans-activating transcription regulator - simian immunodeficiency virus SIVcpz

C;Species: simian immunodeficiency virus SIVcpz

A;Note: host Pan troglodytes (chimpanzee)

C;Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 09-Jul-2004

C;Accession: S09987

R;Huet, T.; Cheynier, R.; Meyers, A.; Roelants, G.; Wain-Hobson, S.

Nature 345, 356-359, 1990

A;Title: Genetic organization of a chimpanzee lentivirus related to HIV-1.

A;Reference number: S09983; MUID:90259077; PMID:2188136

A;Accession: S09987

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-100 &lt;HUE&gt;

A;Cross-references: UNIPROT:P17285; EMBL:X52154; NID:g58866; PIDN:CAA36404.1; PID:g76308

C;Genetics:

A;Gene: tat

A;Introns: 73/2

C;Superfamily: AIDS trans-activating transcription regulator

C;Keywords: AIDS; immunodeficiency; transcription

Query Match 79.3%; Score 46; DB 1; Length 100;

Best Local Similarity 81.8%; Pred. No. 2.1;

Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 YGRKKRQRRR 11

|||||

Db 47 YGRKKRTTTR 57

## RESULT 14

A34356

protamine - Japanese quail

C;Species: Coturnix coturnix japonica (Japanese quail)

C;Date: 15-Jun-1990 #sequence\_revision 15-Jun-1990 #text\_change 09-Jul-2004

C;Accession: A34356

R;Oliva, R.; Goren, R.; Dixon, G.H.

J. Biol. Chem. 264, 17627-17630, 1989

A;Title: Quail (Coturnix japonica) protamine, full-length cDNA sequence, and the function

A;Reference number: A34356; MUID:90036816; PMID:2808336

A;Accession: A34356

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-57 &lt;OLI&gt;

A;Cross-references: UNIPROT:P14402; GB:M30275; NID:g213612; PIDN:AAA49498.1; PID:g213613

C;Superfamily: sperm histone

C;Keywords: DNA binding; nucleus; phosphoprotein

Query Match 74.1%; Score 43; DB 2; Length 57;

Best Local Similarity 70.0%; Pred. No. 3.9;

Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YGRKKRQRRR 10

|||||

Db 47 YGRRRRRRRR 56

## RESULT 15

D38095

T-cell-specific transcription factor 1 splice form D - human

N;Alternate names: transcription factor TCF-1D

C;Species: Homo sapiens (man)

C;Date: 19-May-1994 #sequence\_revision 19-May-1994 #text\_change 16-Aug-2004

C;Accession: D38095; S61877; S61800

R;van de Wetering, M.; Oosterwegel, M.; Holstege, F.; Dooyes, D.; Suikerbuijk, R.; Geuz

J. Biol. Chem. 267, 8530-8536, 1992

A;Title: The human T cell transcription factor-1 gene. Structure, localization, and prom

A;Reference number: A38095; MUID:92235082; PMID:1569101

A;Accession: D38095

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-250 &lt;VAN&gt;

A;Cross-references: GB:X63901

R;Mayer, K.; Wolff, E.; Clevers, H.; Ballhausen, W.G.

submitted to the EMBL Data Library, January 1995

A;Description: The human high mobility group (HMG)-box transcription factor TCF-1: novel

A;Reference number: S61877

A;Accession: S61877

A;Molecule type: mRNA

A;Residues: 1-250 &lt;WAY&gt;

A;Cross-references: EMBL:Z47361; NID:g619881; PIDN:CAA87439.1; PID:g619882

R;Mayer, K.; Wolff, E.; Clevers, H.; Ballhausen, W.G.

Biochim. Biophys. Acta 1263, 169-172, 1995

A;Title: The human high mobility group (HMG)-box transcription factor TCF-1: novel isofo

A;Reference number: S61796; MUID:95367594; PMID:7640309

A;Accession: S61800

A;Status: nucleic acid sequence not shown

A;Molecule type: mRNA

A;Residues: 244-250 &lt;MAW&gt;

A;Cross-references: EMBL:Z47361

C;Genetics:

A;Introns: 32/3; 68/1; 97/2; 137/2; 191/3; 227/3; 243/3

C;Superfamily: HMG box homology

C;Keywords: alternative splicing; DNA binding; transcription factor

P;151-226/Domain: HMG box homology &lt;HMG1&gt;

Query Match 72.4%; Score 42; DB 2; Length 250;

Best Local Similarity 63.6%; Pred. No. 18;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 YGRKKRQRRR 11

|||||

Db 227 YGRKKRRRREK 237

Search completed: March 31, 2005, 00:31:36

Job time : 3.43846 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2005, 00:19:36 ; Search time 6.13461 Seconds  
(without alignments)  
918.211 Million cell updates/sec

Title: US-10-049-822A-12  
Perfect score: 58  
Sequence: 1 YGRKKRQRRR 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt\_03.\*  
1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	58	100.0	58	1 TAT_HV1B5	P04612 human immun
2	58	100.0	64	2 Q6QAV2	Q6QAV2 human immun
3	58	100.0	65	2 Q75540	Q75540 human immun
4	58	100.0	65	2 Q75544	Q75544 human immun
5	58	100.0	68	2 Q6QAV4	Q6QAV4 human immun
6	58	100.0	70	2 Q6MQ1	Q6MQ1 human immun
7	58	100.0	71	2 Q40224	Q40224 human immun
8	58	100.0	71	2 Q40225	Q40225 human immun
9	58	100.0	71	2 Q40226	Q40226 human immun
10	58	100.0	71	2 Q40227	Q40227 human immun
11	58	100.0	71	2 Q40228	Q40228 human immun
12	58	100.0	71	2 Q40231	Q40231 human immun
13	58	100.0	71	2 Q40232	Q40232 human immun
14	58	100.0	71	2 Q40233	Q40233 human immun
15	58	100.0	71	2 Q40234	Q40234 human immun
16	58	100.0	71	2 Q8QDX8	Q8QDX8 human immun
17	58	100.0	71	2 Q8QDX2	Q8QDX2 human immun
18	58	100.0	71	2 Q6MP3	Q6MP3 human immun
19	58	100.0	71	2 Q6MP6	Q6MP6 human immun
20	58	100.0	71	2 Q6MP9	Q6MP9 human immun
21	58	100.0	71	2 Q6MQ3	Q6MQ3 human immun
22	58	100.0	71	2 Q6MQ5	Q6MQ5 human immun
23	58	100.0	71	2 Q6MQ7	Q6MQ7 human immun
24	58	100.0	71	2 Q6MQ9	Q6MQ9 human immun
25	58	100.0	71	2 Q6MR1	Q6MR1 human immun
26	58	100.0	71	2 Q6MR5	Q6MR5 human immun
27	58	100.0	71	2 Q6MR7	Q6MR7 human immun
28	58	100.0	71	2 Q6MR9	Q6MR9 human immun
29	58	100.0	71	2 Q6MS1	Q6MS1 human immun
30	58	100.0	71	2 Q6MS3	Q6MS3 human immun
31	58	100.0	71	2 Q6MS5	Q6MS5 human immun

32	58	100.0	71	2	Q6MS7	human immun
33	58	100.0	71	2	Q6MS9	human immun
34	58	100.0	71	2	Q6MT1	human immun
35	58	100.0	71	2	Q71875	human immun
36	58	100.0	71	2	Q71886	human immun
37	58	100.0	71	2	Q71891	human immun
38	58	100.0	71	2	Q71898	human immun
39	58	100.0	71	2	Q71905	human immun
40	58	100.0	71	2	Q71912	human immun
41	58	100.0	71	2	Q71919	human immun
42	58	100.0	71	2	Q71926	human immun
43	58	100.0	71	2	Q71932	human immun
44	58	100.0	71	2	Q71939	human immun
45	58	100.0	71	2	Q71945	human immun

## ALIGNMENTS

RESULT: 1

TAT\_HV1B5  
ID TAT HV1B5 STANDARD; PRT; 58 AA.  
AC P04612;  
DT 13-AUG-1987 (Rel. 05, Created)  
DT 13-AUG-1987 (Rel. 05, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE TAT protein (Transactivating regulatory protein) (Fragment).  
GN Name=TAT;  
OS Human immunodeficiency virus type 1 (BH5 isolate) (HIV-1).  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11682;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=85111123; PubMed=2578615;  
RA Ratner L., Haseltine W.A., Patarca R., Livak K.J., Starcich B.R.,  
RA Josephs S.F., Doran E.R., Rafalski J.A., Whitehorn E.A.,  
RA Baumanster K., Ivanoff L., Petteway S.R. Jr., Pearson M.D.,  
RA Lautenberger J.A., Papas T.S., Grayeb J., Chang N.T., Gallo R.C.,  
RA Wong-Staal F.;  
RT "Complete nucleotide sequence of the AIDS virus, HTLV-III.";  
RL Nature 313:277-284(1985).  
CC -!- FUNCTION: Transcriptional regulator that acts by binding to the  
trans-activating responsive sequence (TAR) RNA element and  
activates transcription initiation and/or elongation from the LTR  
promoter.  
CC -!- SUBUNIT: Binds cyclin T1 (By similarity).  
CC -!- SUBCELLULAR LOCATION: Nuclear; nucleolar.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; K02012; AAA44656.1; --  
CC HSSP; P04610; 1JFW.  
CC HIV; K02012; TAT\$BH5.  
CC InterPro; IPR001831; IV\_Tat.  
CC Pfam; PF00539; Tat; 1.  
CC PRINTS; PR00055; HIVTATDOMAIN.  
CC Activator; AIDS; Nuclear protein; RNA-binding;  
CC Transcription regulation.  
CC KW NON TER 58 58  
CC FT SEQUENCE 58 AA; 6800 MW; E36C21F8FFD813E3 CRC64;  
CC -----

Query Match 100.0%; Score 58; DB 1; Length 58;  
Best local similarity 100.0%; Pred. No. 0.015;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YGRKKRQRRR 11  
|||||||

```

Db          47 YGRKKRQRRR 57

RESULT 2
Q6QAV2      PRELIMINARY;      PRT;      64 AA.
AC Q6QAV2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Tat protein (Fragment).
GN Name=tat;
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RA Lorenzo E., Herrera R.J., Lai S., Fischl M.A., Hill M.D.;
RA Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
RT "HIV Type 1 Transmission by Human Bite.";
RL AIDS Res. Hum. Retroviruses 20:349-350(2004).
CC -!- FUNCTION: Transcriptional regulator that acts by binding to the
CC trans-activating responsive sequence (TAR) RNA element and
CC activates transcription initiation and/or elongation from the LTR
CC promoter (By similarity).
DR EMBL: AY549944; AAS67918.1; -.
DR GO: GO:0042025; C:host cell nucleus; IEA.
DR GO: GO:0005634; C:nucleus; IEA.
DR GO: GO:0003723; F:RNA binding; IEA.
DR GO: GO:0003700; F:transcription factor activity; IEA.
DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro: IPR001831, IV_Tat.
DR Pfam: PF00539; Tat; 1.
DR PRINTS: PR00055; HIVTATDOMAIN.
DR Activator; Nuclear protein; RNA-binding; Transcription;
DR Transcription regulation.
KW NON TER 1
FT SEQUENCE 64 AA; 7474 MW; E13926D6C42D72D3 CRC64;
SQ

Query Match 100.0%; Score 58; DB 2; Length 64;
Best Local Similarity 100.0%; Pred. No. 0.016;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YGRKKRQRRR 11
| | | | | | | |
Db 39 YGRKKRQRRR 49

RESULT 3
Q75540      PRELIMINARY;      PRT;      65 AA.
AC Q75540;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Tat protein (Fragment).
GN Name=tat;
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RA Lorenzo E., Herrera R.J., Lai S., Fischl M.A., Hill M.D.;
RA Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
RT trans-activating responsive sequence (TAR) RNA element and
CC trans-activating responsive sequence (TAR) RNA element and
CC activates transcription initiation and/or elongation from the LTR
CC promoter (By similarity).
DR EMBL: U57248; AAB17808.1; -.
DR HSP; P12506; ITC.
DR GO: GO:0042025; C:host cell nucleus; IEA.
DR GO: GO:0005634; C:nucleus; IEA.
DR GO: GO:0003723; F:RNA binding; IEA.
DR GO: GO:0003700; F:transcription factor activity; IEA.
DR InterPro: IPR001831, IV_Tat.
DR Pfam: PF00539; Tat; 1.
DR Activator; Nuclear protein; RNA-binding; Transcription;
DR Transcription regulation.
KW NON TER 1
FT SEQUENCE 65 AA; 7677 MW; 91584F861F6A8362 CRC64;
SQ

Query Match 100.0%; Score 58; DB 2; Length 65;
Best Local Similarity 100.0%; Pred. No. 0.017;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YGRKKRQRRR 11
| | | | | | | |
Db 47 YGRKKRQRRR 57

RESULT 4
Q75544      PRELIMINARY;      PRT;      65 AA.
AC Q75544;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Tat protein (Fragment).
GN Name=tat;
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RA Lorenzo E., Herrera R.J., Lai S., Fischl M.A., Hill M.D.;
RA Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
RT trans-activating responsive sequence (TAR) RNA element and
CC trans-activating responsive sequence (TAR) RNA element and
CC activates transcription initiation and/or elongation from the LTR
CC promoter (By similarity).
DR EMBL: U57252; AAB17812.1; -.
DR HSP; P12506; ITC.
DR GO: GO:0042025; C:host cell nucleus; IEA.
DR GO: GO:0005634; C:nucleus; IEA.
DR GO: GO:0003723; F:RNA binding; IEA.
DR GO: GO:0003700; F:transcription factor activity; IEA.
DR InterPro: IPR001831, IV_Tat.
DR Pfam: PF00539; Tat; 1.
DR Activator; Nuclear protein; RNA-binding; Transcription;
DR Transcription regulation.
KW NON TER 1
FT SEQUENCE 65 AA; 7677 MW; 91584F861F6A8362 CRC64;
SQ

Query Match 100.0%; Score 58; DB 2; Length 65;
Best Local Similarity 100.0%; Pred. No. 0.017;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YGRKKRQRRR 11
| | | | | | | |
Db 47 YGRKKRQRRR 57

RESULT 5
Q6QAV4      PRELIMINARY;      PRT;      68 AA.
AC Q6QAV4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Tat protein (Fragment).
GN Name=tat;
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.

```

```

DR GO: GO:0003700; F:transcription factor activity; IEA.
DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro: IPR001831, IV_Tat.
DR Pfam: PF00539; Tat; 1.
DR PRINTS: PR00055; HIVTATDOMAIN.
DR Activator; Nuclear protein; RNA-binding; Transcription;
DR Transcription regulation.
FT NON TER 1
SQ SEQUENCE 65 AA; 7619 MW; 91584F861A2F9736 CRC64;

Query Match 100.0%; Score 58; DB 2; Length 65;
Best Local Similarity 100.0%; Pred. No. 0.017;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YGRKKRQRRR 11
| | | | | | | |
Db 47 YGRKKRQRRR 57

RESULT 4
Q75544      PRELIMINARY;      PRT;      65 AA.
AC Q75544;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Tat protein (Fragment).
GN Name=tat;
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RA Lorenzo E., Herrera R.J., Lai S., Fischl M.A., Hill M.D.;
RA Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
RT trans-activating responsive sequence (TAR) RNA element and
CC trans-activating responsive sequence (TAR) RNA element and
CC activates transcription initiation and/or elongation from the LTR
CC promoter (By similarity).
DR EMBL: U57252; AAB17812.1; -.
DR HSP; P12506; ITC.
DR GO: GO:0042025; C:host cell nucleus; IEA.
DR GO: GO:0005634; C:nucleus; IEA.
DR GO: GO:0003723; F:RNA binding; IEA.
DR GO: GO:0003700; F:transcription factor activity; IEA.
DR InterPro: IPR001831, IV_Tat.
DR Pfam: PF00539; Tat; 1.
DR Activator; Nuclear protein; RNA-binding; Transcription;
DR Transcription regulation.
KW NON TER 1
FT SEQUENCE 65 AA; 7677 MW; 91584F861F6A8362 CRC64;
SQ

Query Match 100.0%; Score 58; DB 2; Length 65;
Best Local Similarity 100.0%; Pred. No. 0.017;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YGRKKRQRRR 11
| | | | | | | |
Db 47 YGRKKRQRRR 57

RESULT 5
Q6QAV4      PRELIMINARY;      PRT;      68 AA.
AC Q6QAV4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Tat protein (Fragment).
GN Name=tat;
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.

```

```
OX NCBI_TaxID=11676;
RN SEQUENCE FROM N.A.
RP Andreu S.M.S., Barra L.A.C., Costa L.J., Sucupira M.C.A.,
RA Souza I.E.L., Diaz R.S.;
RT "HIV Type 1 Transmission by Human Bite.";
RL AIDS Res. Hum. Retroviruses 20:349-350(2004).
CC -1- FUNCTION: Transcriptional regulator that acts by binding to the
CC trans-activating responsive sequence (TAR) RNA element and
CC activates transcription initiation and/or elongation from the LTR
CC promoter (By similarity).
DR EMBL: AY549942; AAS67916.1; -.
DR GO: GO:0042025; C:host cell nucleus; IEA.
DR GO: GO:0005634; C:nucleus; IEA.
DR GO: GO:0003723; F:RNA binding; IEA.
DR GO: GO:0003700; F:transcription factor activity; IEA.
DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro: IPR001831, IV_Tat.
DR Pfam: PF00539; Tat; 1.
DR PRINTS: PR00055; HIVTATDOMAIN.
DR Activator; Nuclear protein; RNA-binding; Transcription;
KW Transcription regulation.
FT NON TER 1
SQ SEQUENCE 68 AA; 7983 MW; 326A05D50078CD29 CRC64;

Query Match 100.0%; Score 58; DB 2; Length 68;
Best Local Similarity 100.0%; Pred. No. 0.017;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YGRKKRQRRR 11
Db 43 YGRKKRQRRR 53

RESULT 6
ID Q66MQ1 PRELIMINARY; PRT; 70 AA.
AC Q66MQ1;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Tat protein (Fragment).
GN Name=tat;
OS Human immunodeficiency virus 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RA Herring B.L., Grant R.M., Delwart E.L.;
RT "No superinfection among seroconcordant couples after well-defined
RL exposure.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Transcriptional regulator that acts by binding to the
CC trans-activating responsive sequence (TAR) RNA element and
CC activates transcription initiation and/or elongation from the LTR
CC promoter (By similarity).
DR EMBL: AY686119; AAU05440.1; -.
DR InterPro: IPR001831, IV_Tat.
DR Pfam: PF00539; Tat; 1.
DR PRINTS: PR00055; HIVTATDOMAIN.
DR Activator; Nuclear protein; RNA-binding; Transcription;
KW Transcription regulation.
FT NON TER 70
SQ SEQUENCE 70 AA; 8096 MW; 28E39B5672863DBB CRC64;

Query Match 100.0%; Score 58; DB 2; Length 70;
Best Local Similarity 100.0%; Pred. No. 0.018;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YGRKKRQRRR 11
Db 47 YGRKKRQRRR 57
```

```
RESULT 7
O40224 PRELIMINARY; PRT; 71 AA.
ID O40224;
AC O40224;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Tat protein (Fragment).
GN Name=tat;
OS Human immunodeficiency virus 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=AD93-A3tat;
RC MEDLINE=97335179; PubMed=9191845; DOI=10.1006/viro.1997.8586;
RA Kirchhoff F., Greenough T.C., Hamacher M., Sullivan J.L.,
RA Desrosiers R.C.;
RT "Activity of human immunodeficiency virus type 1 promoter/TAR regions
RT and tat1 genes derived from individuals with different rates of
RT disease progression.";
RL Virology 232:319-331(1997).
CC -1- FUNCTION: Transcriptional regulator that acts by binding to the
CC trans-activating responsive sequence (TAR) RNA element and
CC activates transcription initiation and/or elongation from the LTR
CC promoter (By similarity).
DR EMBL: AF000522; AAB62521.1; -.
DR HSP; P04610; 1JFW.
DR GO: GO:0042025; C:host cell nucleus; IEA.
DR GO: GO:0005634; C:nucleus; IEA.
DR GO: GO:0003723; F:RNA binding; IEA.
DR GO: GO:0003700; F:transcription factor activity; IEA.
DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro: IPR001831, IV_Tat.
DR Pfam: PF00539; Tat; 1.
DR Activator; Nuclear protein; RNA-binding; Transcription;
KW Transcription regulation.
FT NON TER 71
SQ SEQUENCE 71 AA; 8211 MW; B3031C7AF5EF30E0 CRC64;

Query Match 100.0%; Score 58; DB 2; Length 71;
Best Local Similarity 100.0%; Pred. No. 0.018;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YGRKKRQRRR 11
Db 47 YGRKKRQRRR 57

RESULT 8
O40225 PRELIMINARY; PRT; 71 AA.
ID O40225;
AC O40225;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Tat protein (Fragment).
GN Name=tat;
OS Human immunodeficiency virus 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=BJ93-A3tat;
RC MEDLINE=97335179; PubMed=9191845; DOI=10.1006/viro.1997.8586;
RA Kirchhoff F., Greenough T.C., Hamacher M., Sullivan J.L.,
RA Desrosiers R.C.;
RT "Activity of human immunodeficiency virus type 1 promoter/TAR regions
RT and tat1 genes derived from individuals with different rates of
RT disease progression.";
RL Virology 232:319-331(1997).
CC -1- FUNCTION: Transcriptional regulator that acts by binding to the
```

```

CC trans-activating responsive sequence (TAR) RNA element and
CC activates transcription initiation and/or elongation from the LTR
CC promoter (By similarity).
DR EMBL; AF000523; AAB62522.1; -.
DR HSSP; P12506; ITBC.
DR GO; GO:0042025; C:host cell nucleus; IEA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001831; IV_Tat.
DR Pfam; PF00539; Tat; 1.
DR PRINTS; PR00055; HIVTATDOMAIN.
KW Activator; Nuclear protein; RNA-binding; Transcription;
KW Transcription regulation.
FT NON TER 71
SQ SEQUENCE 71 AA; 8179 MW; CE411588EB96209F CRC64;

Query Match 100.0%; Score 58; DB 2; Length 71;
Best Local Similarity 100.0%; Pred. No. 0.018;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YGRKKRQRRR 11
Dd |||||
47 YGRKKRQRRR 57

RESULT 9
O40226 PRELIMINARY; PRT; 71 AA.
AC O40226;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Tat protein (Fragment).
GN Name=tat;
OS Human immunodeficiency virus 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BT94-A2tat;
RX MEDLINE=97335179; PubMed=9191845; DOI=10.1006/viro.1997.8586;
RA Kirchhoff F., Greenough T.C., Hamacher M., Sullivan J.L.,
RA Desrosiers R.C.;
RT "Activity of human immunodeficiency virus type 1 promoter/TAR regions
RT and tat1 genes derived from individuals with different rates of
RT disease progression.";
RL Virology 232:319-331(1997).
CC -!- FUNCTION: Transcriptional regulator that acts by binding to the
CC trans-activating responsive sequence (TAR) RNA element and
CC activates transcription initiation and/or elongation from the LTR
CC promoter (By similarity).
DR EMBL; AF000523; AAB62524.1; -.
DR HSSP; P04610; IJFW.
DR GO; GO:0042025; C:host cell nucleus; IEA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001831; IV_Tat.
DR Pfam; PF00539; Tat; 1.
DR PRINTS; PR00055; HIVTATDOMAIN.
KW Activator; Nuclear protein; RNA-binding; Transcription;
KW Transcription regulation.
FT NON TER 71
SQ SEQUENCE 71 AA; 8247 MW; C880BD85C90230E5 CRC64;

Query Match 100.0%; Score 58; DB 2; Length 71;
Best Local Similarity 100.0%; Pred. No. 0.018;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YGRKKRQRRR 11
Dd |||||
47 YGRKKRQRRR 57

RESULT 11
O40228 PRELIMINARY; PRT; 71 AA.
AC O40228;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Tat protein (Fragment).
GN Name=tat;
OS Human immunodeficiency virus 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DJ93-A2tat;
RX MEDLINE=97335179; PubMed=9191845; DOI=10.1006/viro.1997.8586;
RA Kirchhoff F., Greenough T.C., Hamacher M., Sullivan J.L.,
RA Desrosiers R.C.;
RT "Activity of human immunodeficiency virus type 1 promoter/TAR regions
RT and tat1 genes derived from individuals with different rates of
RT disease progression.";
RL Virology 232:319-331(1997).
CC -!- FUNCTION: Transcriptional regulator that acts by binding to the
CC trans-activating responsive sequence (TAR) RNA element and
CC activates transcription initiation and/or elongation from the LTR
CC promoter (By similarity).
DR EMBL; AF000524; AAB62523.1; -.
DR HSSP; P04610; IJFW.
DR GO; GO:0042025; C:host cell nucleus; IEA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001831; IV_Tat.
DR Pfam; PF00539; Tat; 1.
DR PRINTS; PR00055; HIVTATDOMAIN.
KW Activator; Nuclear protein; RNA-binding; Transcription;
KW Transcription regulation.
FT NON TER 71
SQ SEQUENCE 71 AA; 8337 MW; 50B96328495E6A2C CRC64;

Query Match 100.0%; Score 58; DB 2; Length 71;
Best Local Similarity 100.0%; Pred. No. 0.018;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YGRKKRQRRR 11

```

```

Db |||||
47 YGRKKRQRRR 57

RESULT 10
O40227 PRELIMINARY; PRT; 71 AA.
AC O40227;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Tat protein (Fragment).
GN Name=tat;
OS Human immunodeficiency virus 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CH94-A2tat;
RX MEDLINE=97335179; PubMed=9191845; DOI=10.1006/viro.1997.8586;
RA Kirchhoff F., Greenough T.C., Hamacher M., Sullivan J.L.,
RA Desrosiers R.C.;
RT "Activity of human immunodeficiency virus type 1 promoter/TAR regions
RT and tat1 genes derived from individuals with different rates of
RT disease progression.";
RL Virology 232:319-331(1997).
CC -!- FUNCTION: Transcriptional regulator that acts by binding to the
CC trans-activating responsive sequence (TAR) RNA element and
CC activates transcription initiation and/or elongation from the LTR
CC promoter (By similarity).
DR EMBL; AF000525; AAB62524.1; -.
DR HSSP; P04610; IJFW.
DR GO; GO:0042025; C:host cell nucleus; IEA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001831; IV_Tat.
DR Pfam; PF00539; Tat; 1.
DR PRINTS; PR00055; HIVTATDOMAIN.
KW Activator; Nuclear protein; RNA-binding; Transcription;
KW Transcription regulation.
FT NON TER 71
SQ SEQUENCE 71 AA; 8247 MW; C880BD85C90230E5 CRC64;

Query Match 100.0%; Score 58; DB 2; Length 71;
Best Local Similarity 100.0%; Pred. No. 0.018;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YGRKKRQRRR 11
Dd |||||
47 YGRKKRQRRR 57

RESULT 11
O40228 PRELIMINARY; PRT; 71 AA.
AC O40228;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Tat protein (Fragment).
GN Name=tat;
OS Human immunodeficiency virus 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DJ93-A2tat;
RX MEDLINE=97335179; PubMed=9191845; DOI=10.1006/viro.1997.8586;
RA Kirchhoff F., Greenough T.C., Hamacher M., Sullivan J.L.,
RA Desrosiers R.C.;
RT "Activity of human immunodeficiency virus type 1 promoter/TAR regions

```

RT and tat1 genes derived from individuals with different rates of  
 RT disease progression.";  
 RL Virology 232:319-331(1997).  
 CC -1- FUNCTION: Transcriptional regulator that acts by binding to the  
 CC trans-activating responsive sequence (TAR) RNA element and  
 CC activates transcription initiation and/or elongation from the LTR  
 CC promoter (By similarity).  
 DR EMBL; AF000526; AAB62525.1; -.  
 DR HSP; P04610; IJFW.  
 DR GO; GO:0042025; C:host cell nucleus; IEA.  
 DR GO; GO:0005634; C:nucleus; IEA.  
 DR GO; GO:0003723; F:RNA binding; IEA.  
 DR GO; GO:0003700; F:transcription factor activity; IEA.  
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
 DR InterPro; IPR001831; IV\_Tat.  
 DR Pfam; PF00539; Tat; 1.  
 DR PRINTS; PR00055; HIVTATDOMAIN.  
 KW Activator; Nuclear protein; RNA-binding; Transcription;  
 KW Transcription regulation.  
 FT NON TER 71  
 SQ SEQUENCE 71 AA; 8191 MW; C880A34AEAFF30E5 CRC64;  
 Query Match 100.0%; Score 58; DB 2; Length 71;  
 Best Local Similarity 100.0%; Pred. No. 0.018;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 YGRKKRRQRRR 11  
 DB 47 YGRKKRRQRRR 57  
 RESULT 12  
 O40231 PRELIMINARY; PRT; 71 AA.  
 AC O40231;  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Tat protein (Fragment).  
 GN Name=tat;  
 OS Human immunodeficiency virus 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=HP91-Altat;  
 RX MEDLINE=97335179; PubMed=9191845; DOI=10.1006/viro.1997.8586;  
 RA Kirchhoff F., Greenough T.C., Hamacher M., Sullivan J.L.,  
 RA Desrosiers R.C.;  
 RT "Activity of human immunodeficiency virus type 1 promoter/TAR regions  
 RT and tat1 genes derived from individuals with different rates of  
 RT disease progression.";  
 RL Virology 232:319-331(1997).  
 CC -1- FUNCTION: Transcriptional regulator that acts by binding to the  
 CC trans-activating responsive sequence (TAR) RNA element and  
 CC activates transcription initiation and/or elongation from the LTR  
 CC promoter (By similarity).  
 DR EMBL; AF000530; AAB62529.1; -.  
 DR HSP; P04610; IJFW.  
 DR GO; GO:0042025; C:host cell nucleus; IEA.  
 DR GO; GO:0005634; C:nucleus; IEA.  
 DR GO; GO:0003723; F:RNA binding; IEA.  
 DR GO; GO:0003700; F:transcription factor activity; IEA.  
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
 DR InterPro; IPR000345; CytC\_heme\_BS.  
 DR InterPro; IPR001831; IV\_Tat.  
 DR Pfam; PF00539; Tat; 1.  
 DR PROSITE; PS00190; CYTOCHROME C; UNKNOWN 1.  
 KW Activator; Nuclear protein; RNA-binding; Transcription;  
 KW Transcription regulation.  
 FT NON TER 71  
 SQ SEQUENCE 71 AA; 8226 MW; 324F908AF030E216 CRC64;  
 Query Match 100.0%; Score 58; DB 2; Length 71;  
 Best Local Similarity 100.0%; Pred. No. 0.018;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 YGRKKRRQRRR 11  
 DB 47 YGRKKRRQRRR 57  
 RESULT 14  
 O40233 PRELIMINARY; PRT; 71 AA.  
 AC O40233;  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Tat protein (Fragment).  
 GN Name=tat;  
 OS Human immunodeficiency virus 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=HP93-A3tat;  
 RX MEDLINE=97335179; PubMed=9191845; DOI=10.1006/viro.1997.8586;  
 RA Kirchhoff F., Greenough T.C., Hamacher M., Sullivan J.L.,  
 RA Desrosiers R.C.;  
 RT "Activity of human immunodeficiency virus type 1 promoter/TAR regions  
 RT and tat1 genes derived from individuals with different rates of  
 RT disease progression.";  
 RL Virology 232:319-331(1997).  
 CC -1- FUNCTION: Transcriptional regulator that acts by binding to the  
 CC trans-activating responsive sequence (TAR) RNA element and  
 CC activates transcription initiation and/or elongation from the LTR  
 CC promoter (By similarity).  
 DR EMBL; AF000529; AAB62528.1; -.  
 DR HSP; P04610; IJFW.  
 DR GO; GO:0042025; C:host cell nucleus; IEA.  
 DR GO; GO:0005634; C:nucleus; IEA.  
 DR GO; GO:0003723; F:RNA binding; IEA.  
 DR GO; GO:0003700; F:transcription factor activity; IEA.  
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
 DR InterPro; IPR001831; IV\_Tat.  
 DR Pfam; PF00539; Tat; 1.  
 DR PRINTS; PR00055; HIVTATDOMAIN.  
 KW Activator; Nuclear protein; RNA-binding; Transcription;  
 KW Transcription regulation.  
 FT NON TER 71  
 SQ SEQUENCE 71 AA; 8162 MW; 5AFE3242E82321B3 CRC64;  
 Query Match 100.0%; Score 58; DB 2; Length 71;

Best Local Similarity 100.0%; Pred. No. 0.018;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 YGRKKRRQRRR 11  
 DB 47 YGRKKRRQRRR 57  
 RESULT 13  
 O40232 PRELIMINARY; PRT; 71 AA.  
 ID O40232;  
 AC O40232;  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Tat protein (Fragment).  
 GN Name=tat;  
 OS Human immunodeficiency virus 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=HP91-Altat;  
 RX MEDLINE=97335179; PubMed=9191845; DOI=10.1006/viro.1997.8586;  
 RA Kirchhoff F., Greenough T.C., Hamacher M., Sullivan J.L.,  
 RA Desrosiers R.C.;  
 RT "Activity of human immunodeficiency virus type 1 promoter/TAR regions  
 RT and tat1 genes derived from individuals with different rates of  
 RT disease progression.";  
 RL Virology 232:319-331(1997).  
 CC -1- FUNCTION: Transcriptional regulator that acts by binding to the  
 CC trans-activating responsive sequence (TAR) RNA element and  
 CC activates transcription initiation and/or elongation from the LTR  
 CC promoter (By similarity).  
 DR EMBL; AF000530; AAB62529.1; -.  
 DR HSP; P04610; IJFW.  
 DR GO; GO:0042025; C:host cell nucleus; IEA.  
 DR GO; GO:0005634; C:nucleus; IEA.  
 DR GO; GO:0003723; F:RNA binding; IEA.  
 DR GO; GO:0003700; F:transcription factor activity; IEA.  
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
 DR InterPro; IPR000345; CytC\_heme\_BS.  
 DR InterPro; IPR001831; IV\_Tat.  
 DR Pfam; PF00539; Tat; 1.  
 DR PROSITE; PS00190; CYTOCHROME C; UNKNOWN 1.  
 KW Activator; Nuclear protein; RNA-binding; Transcription;  
 KW Transcription regulation.  
 FT NON TER 71  
 SQ SEQUENCE 71 AA; 8226 MW; 324F908AF030E216 CRC64;  
 Query Match 100.0%; Score 58; DB 2; Length 71;  
 Best Local Similarity 100.0%; Pred. No. 0.018;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 YGRKKRRQRRR 11  
 DB 47 YGRKKRRQRRR 57  
 RESULT 14  
 O40233 PRELIMINARY; PRT; 71 AA.  
 ID O40233;  
 AC O40233;  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Tat protein (Fragment).  
 GN Name=tat;  
 OS Human immunodeficiency virus 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN=MR94-A3tat;  
RA MEDLINE=97335179; PubMed=9191845; DOI=10.1006/viro.1997.8586;  
RX Kirchhoff F., Greenough T.C., Hamacher M., Sullivan J.L.,  
RA Desrosiers R.C.;  
RT "Activity of human immunodeficiency virus type 1 promoter/TAR regions  
RT and tat1 genes derived from individuals with different rates of  
RT disease progression.";  
RL Virology 232:319-331(1997).  
CC -!- FUNCTION: Transcriptional regulator that acts by binding to the  
CC trans-activating responsive sequence (TAR) RNA element and  
CC activates transcription initiation and/or elongation from the LTR  
CC promoter (By similarity).  
DR EMBL; AF000531; AAB62530.1; -.  
DR HSSP; P12506; 1TBC.  
DR GO; GO:0042025; C:host cell nucleus; IEA.  
DR GO; GO:0005634; C:nucleus; IEA.  
DR GO; GO:0003723; F:RNA binding; IEA.  
DR GO; GO:0003700; F:transcription factor activity; IEA.  
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
DR InterPro; IPR001831; IV\_Tat.  
DR Pfam; PF00539; Tat; 1.  
KW Activator; Nuclear protein; RNA-binding; Transcription;  
KW Transcription regulation.  
FT NON\_TER 71  
SQ SEQUENCE 71 AA; 8034 MW; 5BBA79FB327EA3CD CRC64;  
  
Query Match 100.0%; Score 58; DB 2; Length 71;  
Best Local Similarity 100.0%; Pred. No. 0.018;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 YGRKKRQRRR 11  
|||||  
DB 47 YGRKKRQRRR 57  
  
RESULT 15  
O40234  
ID O40234 PRELIMINARY; PRT; 71 AA.  
AC O40234;  
DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Tat protein (Fragment).  
GN Name=tat;  
OS Human immunodeficiency virus 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=PC93-A2tat;  
RX MEDLINE=97335179; PubMed=9191845; DOI=10.1006/viro.1997.8586;  
RA Kirchhoff F., Greenough T.C., Hamacher M., Sullivan J.L.,  
RA Desrosiers R.C.;  
RT "Activity of human immunodeficiency virus type 1 promoter/TAR regions  
RT and tat1 genes derived from individuals with different rates of  
RT disease progression.";  
RL Virology 232:319-331(1997).  
CC -!- FUNCTION: Transcriptional regulator that acts by binding to the  
CC trans-activating responsive sequence (TAR) RNA element and  
CC activates transcription initiation and/or elongation from the LTR  
CC promoter (By similarity).  
DR EMBL; AF000532; AAB62531.1; -.  
DR HSSP; P04610; 1JFW.  
DR GO; GO:0042025; C:host cell nucleus; IEA.  
DR GO; GO:0005634; C:nucleus; IEA.  
DR GO; GO:0003723; F:RNA binding; IEA.  
DR GO; GO:0003700; F:transcription factor activity; IEA.  
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
DR InterPro; IPR001831; IV\_Tat.  
DR Pfam; PF00539; Tat; 1.  
KW Activator; Nuclear protein; RNA-binding; Transcription;  
KW Transcription regulation.  
FT NON\_TER 71

SQ SEQUENCE 71 AA; 8235 MW; A8030EF1E71A863D CRC64;  
  
Query Match 100.0%; Score 58; DB 2; Length 71;  
Best Local Similarity 100.0%; Pred. No. 0.018;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 YGRKKRQRRR 11  
|||||  
DB 47 YGRKKRQRRR 57  
  
Search completed: March 31, 2005, 00:30:52  
Job time : 8.13461 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2005, 00:17:06 ; Search time 11.2 Seconds  
(without alignments)  
552.515 Million cell updates/sec

Title: US-10-049-822A-13  
Perfect score: 93  
Sequence: 1 RQIKWIFQNRMKWK 16

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*  
1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	93	100.0	16	4	AAB73306 Drosophil
2	93	100.0	16	7	ADF56535 Homeobox
3	85	91.4	17	3	AAY83570 Peptide f
4	84	90.3	16	2	AAW45974 Cysteine
5	84	90.3	16	2	AAW33407 Peptide 4
6	84	90.3	16	2	AAW33410 D-form pe
7	84	90.3	16	2	AAW82958 Oestrogen
8	84	90.3	16	2	AAW56397 Preferred
9	84	90.3	16	2	AAW71270 Antennape
10	84	90.3	16	2	AAW71316 Antennape
11	84	90.3	16	2	AAW30508 Drosophil
12	84	90.3	16	2	AAW91046 Internali
13	84	90.3	16	2	AAW52102 Peptide f
14	84	90.3	16	2	AAW00859 Peptide p
15	84	90.3	16	2	AAW13509 Signal se
16	84	90.3	16	3	AAW87920 Drosophil
17	84	90.3	16	3	AAW27060 Beta-cate
18	84	90.3	16	3	AAW93667 Peptide w
19	84	90.3	16	3	AAW67966 Carboxyl
20	84	90.3	16	3	AAW93551 Amino aci
21	84	90.3	16	3	AAW55818 Signal se
22	84	90.3	16	3	AAW71008 Drosophil
23	84	90.3	16	3	AAW51212 Antennape
24	84	90.3	16	3	AAW51167 Drosophil
25	84	90.3	16	3	AAW10343 Peptide A

26	84	90.3	16	3	AAB19251 Fragment
27	84	90.3	16	3	AAY93178 Protegrin
28	84	90.3	16	3	AAB35694 Peptide a
29	84	90.3	16	3	AAB22025 Membrane
30	84	90.3	16	3	AAB29423 ANTP pept
31	84	90.3	16	3	AAB03927 Internali
32	84	90.3	16	3	AAY93954 Peptide u
33	84	90.3	16	3	AAB29574 Antennape
34	84	90.3	16	3	Adel14785 Carrier m
35	84	90.3	16	3	Adel14761 Drosophil
36	84	90.3	16	4	AAB73091 Rheumatoi
37	84	90.3	16	4	AAB60004 Internali
38	84	90.3	16	4	AAB70753 Cell memb
39	84	90.3	16	4	Aae02974 Protein t
40	84	90.3	16	4	AAB60671 Antennape
41	84	90.3	16	4	AAU06064 Drosophil
42	84	90.3	16	4	AAB49914 HIF-1alph
43	84	90.3	16	4	AAB66996 Antennape
44	84	90.3	16	4	AAU00813 Fruit fly
45	84	90.3	16	4	AAE12205 Membrane

## ALIGNMENTS

## RESULT.1

AAB73306  
ID AAB73306 standard; protein; 16 AA.

XX AAB73306;

AC AAB73306;

DT 22-MAY-2001 (first entry)

XX Drosophila antennapedia protein transduction domain (PTD).

KW Antennapedia; protein transduction domain; PTD; fusion protein;

KW rat Bcl-xL mutant; Bcl-xFNK; apoptosis inhibitor; membrane permeable;

KW programmed cell death inhibitor; wild-type; antiapoptotic;

KW cell death-associated disease; tissue transplant preservative.

XX Drosophila sp.

XX WO200112807-A1.

XX 22-FEB-2001.

XX 17-AUG-2000; 2000WO-JP005502.

XX 17-AUG-1999; 99JP-00230842.

XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.

XX Ohta S, Asoh S;

XX WPI; 2001-211219/21.

DR Modified cDNA of rat bcl-x gene and encoded protein with membrane permeability to enhance uptake for effective inhibition of cell death

PT e.g. apoptosis, useful in remedies for diseases associated with cell death.

XX Claim 3; Page 51; 56pp; Japanese.

XX The invention relates to a mutant rat Bcl-x protein and the cDNA encoding it. The mutant rat Bcl-x protein (Bcl-xFNK) has the substitutions Y22F, Q26N, and R165K relative to the wild-type Bcl-xL protein. The invention also encompasses recombinant vectors and host cells comprising the modified nucleic acid sequence. The mutant Bcl-x protein is able to permeate the cell membrane, thus enhancing its ability to be taken up into a cell and to act as an inhibitor of apoptosis (programmed cell death). Bcl-xFNK and nucleic acids encoding it are useful in remedies for diseases associated with cell death and in additives for maintaining the stability of transplanted cells and organs. The present invention

CC represents the Drosophila antennapedia protein transduction domain (PTD)  
 CC which is specifically claimed for use in generating Bcl-xFNK fusion  
 CC proteins  
 XX SQ Sequence 16 AA;

Query Match 100.0%; Score 93; DB 4; Length 16;  
 Best Local Similarity 100.0%; Pred. No. 6.7e-07;  
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RQIKWIFQNRRMKWFK 16  
 |||||  
 Db 1 RQIKWIFQNRRMKWFK 16  
 |||||

RESULT 2  
 ID ADF56535 standard; peptide; 16 AA.  
 AC ADF56535;  
 XX  
 DT 12-FEB-2004 (first entry)  
 XX  
 DE Homeobox protein Antennapedia protein transduction domain, PTD, SEQ ID 4.  
 XX  
 KW Human; synovial sarcoma; hBRW/hsNF2 alpha;  
 KW cytoplasmic membrane penetrating peptide; cytostatic;  
 KW protein transduction domain; PTD; homeobox; antennapedia.  
 XX  
 OS Drosophila sp.  
 XX  
 PN JP2003252802-A.  
 XX  
 FD 10-SEP-2003.  
 XX  
 XX 27-FEB-2002; 2002JP-00050894.  
 XX  
 PR 27-FEB-2002; 2002JP-00050894.  
 XX  
 FA (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.  
 XX  
 XX WPI; 2003-857284/80.  
 XX  
 XX Novel drug comprising hBRW/hsNF2 alpha polynucleotide, useful for  
 PT treating human synovial sarcoma.  
 FT  
 XX Disclosure; SEQ ID NO 4; 21pp; Japanese.  
 PS  
 CC The present invention relates to a drug (I) for treating human synovial  
 CC sarcoma by introducing into the cell the fusion protein hBRW/hsNF2  
 CC (alpha) (II; ADF56533), or its coding sequence (ADF56532). (II) comprises  
 CC a cytoplasmic membrane penetrating peptide. (I) is useful for treating  
 CC human synovial sarcoma.  
 XX  
 SQ Sequence 16 AA;

Query Match 100.0%; Score 93; DB 7; Length 16;  
 Best Local Similarity 100.0%; Pred. No. 6.7e-07;  
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RQIKWIFQNRRMKWFK 16  
 |||||  
 Db 1 RQIKWIFQNRRMKWFK 16  
 |||||

RESULT 3  
 ID AAY83570 standard; peptide; 17 AA.  
 XX  
 AC AAY83570;  
 XX  
 DT 29-AUG-2000 (first entry)  
 XX  
 XX

DE Peptide fragment of membrane transport vector penetratin.  
 XX Penetratin; translocation; membrane; drug delivery; therapy; treatment;  
 KW conjugate; vector.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 1  
 FT /label= Beta-alanine  
 FT /note= "Biotinylated"  
 XX  
 PN WO200029427-A2.  
 XX  
 XX 25-MAY-2000.  
 XX  
 XX 11-NOV-1999; 99WO-GB003750.  
 XX  
 PR 13-NOV-1998; 98GB-00025000.  
 PR 13-NOV-1998; 98GB-00025001.  
 PR 04-FEB-1999; 99GB-00002522.  
 PR 04-FEB-1999; 99GB-00002525.  
 PR 22-JUN-1999; 99GB-00014578.  
 XX  
 PA (CYCL-) CYCLACEL LTD.  
 XX  
 FI Fischer MP, Zhelev N;  
 XX  
 DR WPI; 2000-387734/33.  
 XX  
 XX New membrane translocation peptide carrier group for delivering  
 PT therapeutic agents into target cells comprises specified sequence of  
 FT amino acids.  
 FT  
 XX Claim 9; Page 27; 59pp; English.  
 PS  
 CC Penetratin is a membrane translocation polypeptide and as such, active  
 CC peptide fragments of penetratin can be used to translocate conjugated  
 CC molecules across the cell membrane e.g. drugs. 13 chemical entities are  
 CC described for use in the method such as paclitaxel-  
 CC 2'succinimidopropionyl CbetAA-RRMKWK-NH 2, and podophyllotoxin-4-  
 CC succinimidopropionyl CbetAA-RRMKWK-NH 2. The method has applications as  
 CC a drug delivery system for treatment and therapy. The resulting  
 CC conjugated molecules exhibit high immunogenicity, solubility and  
 CC clearance. The penetratin peptide fragment may be truncated and or have  
 CC amino acid substitutions. See GENESEQ records AAY83520-Y83581  
 XX  
 SQ Sequence 17 AA;

Query Match 91.4%; Score 85; DB 3; Length 17;  
 Best Local Similarity 93.8%; Pred. No. 1.1e-05;  
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RQIKWIFQNRRMKWFK 16  
 |||||  
 Db 2 RQIKWIFQNRRMKWAK 17  
 |||||

RESULT 4  
 ID AAW45974 standard; peptide; 16 AA.  
 XX  
 AC AAW45974;  
 XX  
 DT 01-JUL-1998 (first entry)  
 XX  
 DE Cysteine protease inhibiting peptide for preventing cell death.  
 KW Neuronal cell death; neurodegenerative disorder; inhibition;  
 KW cysteine protease; cardiovascular; liver disease.  
 XX  
 OS Synthetic.  
 XX





PI Chassaing G, Prochiantz A;  
 XX WPI; 1997-226166/20.  
 XX  
 XX New peptide(s) of high hydrophobic amino acid content - useful as vectors  
 PT for delivering peptides and nucleic acids to cells.  
 XX  
 XX Disclosure; Page 7; 35pp; French.  
 XX  
 XX New peptides are provided which are 16 amino acids long and which are  
 CC analogues of the peptide corresponding to residues 43-58 of the  
 CC Antennapedia transcription factor homeodomain (AntpHD). The peptides  
 CC contain 6-10 hydrophobic amino acids. They have the general formula: X1-  
 CC X2-X3-X4-X5-Trp-X7-X8-X9-X10-X11-X12-X13-X14-X15-X16-X17-X18-X19-X20-  
 CC X12-X11-X10-X9-X8-X7-Trp-X5-X4-X3-X2-X1 in which X1-X5 and X7-X16 are any  
 CC alpha-amino acids, provided that: (1) the peptide contains 6-10  
 CC hydrophobic amino acids; (2) X3 and X5 are not both Val; and (3) the  
 CC natural Antp 43-58 sequence RQIKWFWQRMRMKWK (see AAW33407) is excluded.  
 CC The present sequence (the D-form of the 43-58 peptide) is a specific  
 CC example of the new peptides. The peptides are used as vectors for  
 CC introducing into live cells compounds which affect cell function,  
 CC especially peptides and nucleic acids. They can cross cellular membranes  
 CC and reach various cell compartments. They are as effective as helix 3 of  
 CC a homeodomain peptide  
 XX  
 XX Sequence 16 AA;  
 SQ  
 Query Match 90.3%; Score 84; DB 2; Length 16;  
 Best Local Similarity 93.8%; Pred. No. 1.4e-05;  
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 RQIKWFWQRMRMKWFK 16  
 DB 1 RQIKWFWQRMRMKWK 16  
 RESULT 7  
 AAW82958  
 ID AAW82958 standard; peptide; 16 AA.  
 XX  
 XX AAW82958;  
 XX  
 XX 04-FEB-1999 (first entry)  
 XX  
 XX Oestrogen receptor activity inhibiting peptide #14.  
 XX  
 XX Human; oestrogen receptor activity inhibitor; anti-oestrogen; diagnosis;  
 KW breast cancer; estrogen; tumour; phosphotyrosyl peptide;  
 KW malonyltyrosyl peptide; steroid receptor co-activator-1.  
 XX  
 XX Synthetic.  
 OS Homo sapiens.  
 XX  
 XX WO9846250-A1.  
 XX  
 XX 22-OCT-1998.  
 XX  
 XX 14-APR-1998; 98WO-US007711.  
 XX  
 XX 14-APR-1997; 97US-0043545P.  
 XX  
 XX (REGC ) UNIV CALIFORNIA.  
 XX  
 XX Pietras RJ;  
 XX  
 XX WPI; 1998-594522/50.  
 XX  
 XX New anti-oestrogen peptide compositions - comprise sequences based on  
 PT oestrogen receptor and steroid receptor co-activator-1 sequences, used  
 PT for treating cancers.  
 XX  
 XX Claim 59; Page 156; 182pp; English.  
 XX  
 CC The present invention describes a composition comprising an isolated  
 CC oestrogen receptor activity inhibiting (anti-oestrogen) peptide. The  
 CC peptides used in the composition comprise sequences of human oestrogen  
 CC receptor (OR) surrounding Tyr537 and steroid receptor co-activator-1 (SRC  
 CC -1). The peptide compositions, nucleic acids and vectors of the present  
 CC invention can reduce OR activity in a cell, reduced OR polypeptide  
 CC dimerisation in a cell and reduce the binding of SRC-1 polypeptide to an  
 CC OR polypeptide dimer in a cell. They can be used for killing cancer cells  
 CC and treating cancers, particularly breast cancer. The present sequence  
 CC represents a specifically claimed anti-oestrogen peptide  
 XX  
 XX Sequence 16 AA;  
 SQ  
 Query Match 90.3%; Score 84; DB 2; Length 16;  
 Best Local Similarity 93.8%; Pred. No. 1.4e-05;  
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 RQIKWFWQRMRMKWFK 16  
 DB 1 RQIKWFWQRMRMKWK 16  
 RESULT 8  
 AAW56397  
 ID AAW56397 standard; peptide; 16 AA.  
 XX  
 XX AAW56397;  
 XX  
 XX 05-AUG-1998 (first entry)  
 XX  
 XX Preferred signal sequence of the invention.  
 XX  
 XX Signal peptide; nuclear localisation signal; NLS;  
 KW immunosuppressive activity; inhibition; nuclear translocation inhibitor;  
 KW nuclear translocation; treatment; immune disorder; autoimmune disease;  
 KW hypersensitivity; sepsis; prevention; septic shock; antiviral agent;  
 KW tumour growth suppressor.  
 XX  
 XX Unidentified.  
 OS  
 XX WO9811907-A1.  
 XX  
 XX 26-MAR-1998.  
 XX  
 XX 15-SEP-1997; 97WO-US016217.  
 XX  
 XX 20-SEP-1996; 96US-0026978P.  
 XX  
 XX 12-SEP-1997; 97US-00928958.  
 XX  
 XX (BRIM ) BRISTOL-MYERS SQUIBB CO.  
 XX  
 XX Nadler SG, Cleaveland JS, Blake J, Haffar OK;  
 FI WPI; 1998-217028/19.  
 XX  
 XX Nuclear translocation inhibitor polypeptides - comprising signal sequence  
 PT for delivery through the cytoplasmic membrane and at least 2 nuclear  
 PT localisation sequences.  
 XX  
 XX Claim 5; Page 43; 69pp; English.  
 XX  
 XX Peptides AAW56397-99 represent preferred signal sequences of the  
 CC invention. They are used to construct the nuclear translocation inhibitor  
 CC polypeptides of the invention. Nuclear translocation inhibitor  
 CC polypeptides comprise a signal sequence peptide capable of delivering the  
 CC polypeptide through the cytoplasmic membrane into a cell, and at least 2  
 CC nuclear localisation signals (NLSs). The polypeptides can be used to  
 CC inhibit nuclear translocation of a cellular protein. In addition, since  
 CC the nuclear translocation of certain cellular peptides is required for  
 CC the host organism to mount an immune response, the polypeptide inhibitors  
 CC are useful as immunosuppression agents. The polypeptides can therefore be  
 CC used for the treatment of immune disorders including autoimmune diseases.  
 CC The polypeptides can also be used for treating physical symptoms

CC manifested by responses to allergens which can initiate a state of  
 CC hypersensitivity, for the treatment of sepsis and in the prevention of  
 CC septic shock, antiviral agents, tumour growth suppressors, and for  
 CC transcriptionally modulating the expression of cellular genes  
 XX  
 SQ Sequence 16 AA;

Query Match 90.3%; Score 84; DB 2; Length 16;  
 Best Local Similarity 93.8%; Pred. No. 1.4e-05;  
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ROIKIWFONRRMKWFK 16  
 |||||  
 DB 1 ROIKIWFONRRMKWKK 16  
 |||||

RESULT 9  
 AAW71270  
 ID AAW71270 standard; protein; 16 AA.  
 XX  
 AC AAW71270;  
 XX  
 DT 23-NOV-1998 (first entry)  
 XX  
 DE Antennapedia peptide for directing antisense oligonucleotides to a cell.  
 XX  
 KW Antisense oligonucleotide; inhibit; cell death; Nedd2; human Ich-1L gene;  
 KW neuronal cell death; treatment; aging; amyotrophic lateral sclerosis;  
 KW Alzheimer's disease; dentatorubral; pallidolysial atrophy;  
 KW Huntington's disease; Machado-Joseph disease; multiple sclerosis;  
 KW muscular dystrophy; Parkinson's disease; senility;  
 KW spinocerebellar ataxia type I; spinobulbar muscular atrophy; stroke;  
 KW trauma; antennapedia.  
 XX  
 OS Synthetic.  
 XX

Key Location/Qualifiers  
 FH Modified-site 1  
 FT /note= "N-terminal protected with 3-nitro-2-pyridyl  
 FT sulphenyl group (Npys)"  
 FT  
 XX WO9838861-A1.  
 XX  
 PD 11-SEP-1998.  
 XX  
 PF 03-MAR-1998; 98WO-US004128.  
 XX  
 PR 03-MAR-1997; 97US-00810540.  
 XX  
 PA (UYCO ) UNIV COLUMBIA NEW YORK.  
 XX  
 PI Troy CM, Shelanski ML;  
 XX WPI; 1998-506333/43.  
 XX  
 PT Anti-sense oligonucleotide(s) to cysteine aspartase genes - used to  
 PT inhibit neurodegenerative disorder associated with e.g. ageing,  
 PT Alzheimer's, Huntington's or Parkinson's disease.  
 XX  
 PS Claim 5; Page 39; 60pp; English.  
 XX

AAW71270 and AAW71315-16 represent antennapedia peptides which are used  
 to, direct the antisense oligonucleotides (AAV54973-74) of the invention  
 to a cell. The antisense oligonucleotides are used to inhibit cell death  
 mediated by withdrawal of a trophic factor. AAV54973 inhibits the  
 expression of a Nedd2 rodent gene, while AAV54974 inhibits expression of  
 a human Ich-1L gene. The oligonucleotides are used to inhibit neuronal  
 cell death, especially for treatment of neuronal cell death caused by  
 e.g. aging, amyotrophic lateral sclerosis, Alzheimer's disease,  
 dentatorubral and pallidolysial atrophy, Huntington's disease, Machado-  
 Joseph disease, multiple sclerosis, muscular dystrophy, Parkinson's  
 disease, senility, spinocerebellar ataxia type I, spinobulbar muscular  
 atrophy, stroke or trauma

XX Sequence 16 AA;  
 SQ Query Match 90.3%; Score 84; DB 2; Length 16;  
 Best Local Similarity 93.8%; Pred. No. 1.4e-05;  
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ROIKIWFONRRMKWFK 16  
 |||||  
 DB 1 ROIKIWFONRRMKWKK 16  
 |||||

RESULT 10  
 AAW71316  
 ID AAW71316 standard; protein; 16 AA.  
 XX  
 AC AAW71316;  
 XX  
 DT 23-NOV-1998 (first entry)  
 XX  
 DE Antennapedia peptide for directing antisense oligonucleotides to a cell.  
 XX  
 KW Antisense oligonucleotide; inhibit; cell death; Nedd2; human Ich-1L gene;  
 KW neuronal cell death; treatment; aging; amyotrophic lateral sclerosis;  
 KW Alzheimer's disease; dentatorubral; pallidolysial atrophy;  
 KW Huntington's disease; Machado-Joseph disease; multiple sclerosis;  
 KW muscular dystrophy; Parkinson's disease; senility;  
 KW spinocerebellar ataxia type I; spinobulbar muscular atrophy; stroke;  
 KW trauma; antennapedia.  
 XX  
 OS Synthetic.  
 XX

Key Location/Qualifiers  
 FH Modified-site 1  
 FT /note= "N-terminal protected with 3-nitro-2-pyridyl  
 FT sulphenyl group (Npys)"  
 FT  
 XX WO9838861-A1.  
 XX  
 PD 11-SEP-1998.  
 XX  
 PF 03-MAR-1998; 98WO-US004128.  
 XX  
 PR 03-MAR-1997; 97US-00810540.  
 XX  
 PA (UYCO ) UNIV COLUMBIA NEW YORK.  
 XX  
 PI Troy CM, Shelanski ML;  
 XX WPI; 1998-506333/43.  
 XX  
 PT Anti-sense oligonucleotide(s) to cysteine aspartase genes - used to  
 PT inhibit neurodegenerative disorder associated with e.g. ageing,  
 PT Alzheimer's, Huntington's or Parkinson's disease.  
 XX  
 PS Disclosure; Page 14; 60pp; English.  
 XX

AAW71270 and AAW71315-16 represent antennapedia peptides which are used  
 to, direct the antisense oligonucleotides (AAV54973-74) of the invention  
 to a cell. The antisense oligonucleotides are used to inhibit cell death  
 mediated by withdrawal of a trophic factor. AAV54973 inhibits the  
 expression of a Nedd2 rodent gene, while AAV54974 inhibits expression of  
 a human Ich-1L gene. The oligonucleotides are used to inhibit neuronal  
 cell death, especially for treatment of neuronal cell death caused by  
 e.g. aging, amyotrophic lateral sclerosis, Alzheimer's disease,  
 dentatorubral and pallidolysial atrophy, Huntington's disease, Machado-  
 Joseph disease, multiple sclerosis, muscular dystrophy, Parkinson's  
 disease, senility, spinocerebellar ataxia type I, spinobulbar muscular  
 atrophy, stroke or trauma

Query Match 90.3%; Score 84; DB 2; Length 16;

```

Best Local Similarity 93.8%; Pred. NO. 1.4e-05;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RQIKIWQNRRMKWFK 16
Db 1 RQIKIWQNRRMKWKK 16

RESULT 11
AAW30508
ID AAW30508 standard; peptide; 16 AA.
XX
AC AAW30508;
XX
DT 26-OCT-1998 (first entry)
XX
DE Drosophila membrane translocation sequence.
XX
DE DP-1; transcription factor; antagonist; E2F protein; apoptosis;
KW cell proliferation; cardiovascular cell; restenosis; tumor;
KW surgical stent; therapy; membrane translocation; antennapedia protein.
XX
XX Drosophila melanogaster.
OS
XX
XX WO9828334-A1.
XX
XX 02-JUL-1998.
XX
XX 22-DEC-1997; 97WO-GB003506.
XX
XX 20-DEC-1996; 96GB-00026589.
XX
XX (PROL-) PROLIFIX LTD.
XX
XX La Thangue NB, Bandara LR;
XX
XX WPI; 1998-377596/32.
XX
XX Polypeptide fragments of the DP-1 transcription factor - used for
XX inducing apoptosis, specifically in tumour and cardiovascular cells, e.g.
XX for preventing re-stenosis.
XX
XX Disclosure; Page 5; 55pp; English.
XX
XX This polypeptide comprises a membrane translocation sequence derived from
XX the Drosophila melanogaster antennapedia protein. Such membrane
XX translocation sequences are useful in directing entry of a polypeptide
XX into a cell. Polypeptides of the invention (see AAW30504-07) are derived
XX from the DEF box region (see AAW30501) of transcription factor DP1. They
XX act as antagonists of the heterodimerisation of a Dp protein with an E2F
XX protein, and can be used to induce apoptosis, specifically in tumour and
XX cardiovascular cells, e.g. for preventing restenosis. A claimed fusion
XX protein comprises a DEF box peptide and the Drosophila antennapedia
XX protein membrane translocation sequence
XX
SQ Sequence 16 AA;

Query Match 90.3%; Score 84; DB 2; Length 16;
Best Local Similarity 93.8%; Pred. NO. 1.4e-05;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RQIKIWQNRRMKWFK 16
Db 1 RQIKIWQNRRMKWKK 16

RESULT 12
AAW91046
ID AAW91046 standard; peptide; 16 AA.
XX
AC AAW91046;
XX
DT 24-MAR-1999 (first entry)
XX

Internalization sequence associated with cadherin modulating agents.

Alpha-catenin; beta-catenin; interaction; modulation; cell adhesion;
cadherin-mediated function; demyelinating neurological disease;
multiple sclerosis; drug delivery; cancer; angiogenesis; immune system;
central nervous system; apoptosis induction; cadherin-expression cell;
pregnancy prevention; vasopermeability; synaptic stability; diabetes;
rheumatoid arthritis; allergic response; learning; antennapedia protein;
antibody-mediated graft rejection; internalization sequence; memory.

Synthetic.
WO9845319-A2.
15-OCT-1998.
14-APR-1998; 98WO-CA000322.
10-APR-1997; 97US-0043361P.
(UTMC-) UNIV MCGILL.
Blaschuk OW, Gour BJ;
WPI; 1999-024009/02.
New catenin modulating agents - comprising peptides having a sequence HAV
or analogues or antibodies, used for modulating cadherin-mediated
functions.
Claim 16; Page 78; 106pp; English.

The present sequence represents antennapedia protein derived
internalization sequence associated with cadherin modulating agents.
These agents are used for modulating cadherin-mediated functions. They
can be used for disrupting interaction between alpha-catenin and beta-
catenin in a cell, inhibiting cell adhesion, e.g. between epithelial
cells, endothelial cells, neural cells, tumour cells and lymphocytes, for
treating a demyelinating neurological disease, e.g. multiple sclerosis,
for reducing unwanted cellular adhesion in a mammal, for enhancing the
delivery of a drug through the skin of a mammal, for enhancing the
delivery of a drug to a tumour in a mammal, for treating cancer in a
mammal, for inhibiting angiogenesis in a mammal, for enhancing drug
delivery to the central nervous system of a mammal, for inducing
apoptosis in a cadherin-expression cell, for modulating the immune system
of a mammal, for preventing pregnancy in a mammal, for increasing
vasopermeability in a mammal, or for inhibiting synaptic stability in a
mammal. In particular they can be used for treating diabetes, rheumatoid
arthritis, allergic responses, antibody-mediated graft rejection or for
stimulating learning and memory
Sequence 16 AA;

Query Match 90.3%; Score 84; DB 2; Length 16;
Best Local Similarity 93.8%; Pred. NO. 1.4e-05;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RQIKIWQNRRMKWFK 16
Db 1 RQIKIWQNRRMKWKK 16

RESULT 13
AAW52102
ID AAW52102 standard; peptide; 16 AA.
XX
AC AAW52102;
XX
DT 28-JAN-2000 (first entry)
XX
XX Peptide from the third helix of antennapedia homeodomain protein.

```



PA (NEW-) NEW ENGLAND MEDICAL CENT HOSPITALS INC.

XX PI Potter DA, Skolnik PR;

XX DR WPI; 1999-326923/27.

XX FT Fusion protein of signal sequence and calpastatin.

XX PS Disclosure; Page 5; 46pp; English.

XX CC The invention features fusion proteins that contain a calpastatin peptide  
 CC and a signal sequence capable of delivering the fusion protein into a  
 CC eukaryotic cell. The fusion protein is used for the inhibition of calpain  
 CC in a cell. The fusion proteins are specifically used: (a) to prevent  
 CC aggregation and degranulation of platelets (e.g. during storage); (b) to  
 CC inhibit hypoxia-induced sickling of erythrocytes (during storage,  
 CC facilitating subsequent transfusion of autologous cells for treatment of  
 CC sickle cell crises); and (c) to inhibit activation of human immune  
 CC deficiency virus provirus in infected cells (or similarly for other  
 CC viruses regulated by NF-kappaB). Other disclosed uses are: to treat or  
 CC prevent inflammation (e.g. arthritis or asthma), unwanted immune  
 CC responses (e.g. transplant rejection), restenosis (associated with  
 CC angioplasty), cancer, subarachnoid hemorrhage, vasospasm, muscular  
 CC dystrophy, cataracts and traumatic birth injury; to prevent spread of  
 CC platelets on surfaces (e.g. when applied to the surface of stents,  
 CC catheters etc.); to reduce coronary thrombosis in by-pass surgery and  
 CC angioplasty; to treat myocardial infarction, or to prevent progression of  
 CC infarction (myocardial or cerebral). The fusion protein has a reversible  
 CC inhibitory effect and enters cells easily. It allows platelets to be  
 CC stored cold with reduced change in shape

XX SQ Sequence 16 AA;

Query Match 90.3%; Score 84; DB 2; Length 16;  
 Best Local Similarity 93.8%; Pred. No. 1.4e-05;  
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ROIKIWFQNRMRKWK 16  
 |||||  
 Db 1 ROIKIWFQNRMRKWK 16

Search completed: March 31, 2005, 00:28:25  
 Job time : 15.2 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2005, 00:23:16 ; Search time 3.2 Seconds  
(without alignments)  
373.245 Million cell updates/sec

Title: US-10-049-822a-13

Perfect score: 93

Sequence: 1 RQKIWFQRRMKWFK 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*

1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCJTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	84	90.3	16	2	US-08-928-958-7
2	84	90.3	16	2	US-08-810-540-3
3	84	90.3	16	2	US-08-810-540-6
4	84	90.3	16	2	US-09-072-429-7
5	84	90.3	16	3	US-08-964-302A-6
6	84	90.3	16	3	US-09-116-294-4
7	84	90.3	16	3	US-08-964-614A-4
8	84	90.3	16	3	US-08-849-486-1
9	84	90.3	16	3	US-08-849-486-4
10	84	90.3	16	3	US-09-208-966-54
11	84	90.3	16	3	US-09-308-935-8
12	84	90.3	16	3	US-09-441-416A-6
13	84	90.3	16	3	US-09-296-089-33
14	84	90.3	16	3	US-09-419-826-35
15	84	90.3	16	3	US-09-302-305C-10
16	84	90.3	16	4	US-09-346-847-1
17	84	90.3	16	4	US-09-346-847-25
18	84	90.3	16	4	US-09-057-363C-47
19	84	90.3	16	4	US-09-043-560B-3
20	84	90.3	16	4	US-09-648-400A-29
21	84	90.3	16	4	US-09-227-652B-4
22	84	90.3	16	4	US-09-780-070-38
23	84	90.3	16	4	US-08-610-220B-9
24	84	90.3	16	4	US-09-775-052A-54
25	84	90.3	16	4	US-09-155-165-22
26	84	90.3	16	4	US-09-792-480-29
27	84	90.3	16	4	US-09-551-976-33

Sequence 47, Appl  
Sequence 2, Appl  
Sequence 9, Appl  
Sequence 4, Appl  
Sequence 29, Appl  
Sequence 5, Appl  
Sequence 21, Appl  
Sequence 6, Appl  
Sequence 10, Appl  
Sequence 9, Appl  
Sequence 3, Appl  
Sequence 1, Appl  
Sequence 1, Appl  
Sequence 17, Appl  
Sequence 20, Appl  
Sequence 22, Appl  
Sequence 27, Appl  
Sequence 30, Appl

28 84 90.3 16 4 US-09-265-107-47  
29 84 90.3 16 4 US-09-707-263A-2  
30 84 90.3 16 4 US-09-545-433-9  
31 84 90.3 16 4 US-09-720-003C-4  
32 84 90.3 16 4 US-10-209-421-29  
33 84 90.3 16 4 US-09-512-260A-5  
34 84 90.3 16 4 US-09-937-837-21  
35 84 90.3 16 4 US-10-009-049-6  
36 84 90.3 16 4 US-09-959-873B-10  
37 84 90.3 16 4 US-09-150-623-9  
38 84 90.3 16 4 US-09-748-063-3  
39 84 90.3 16 4 US-10-031-505-1  
40 84 90.3 16 4 US-10-144-549-1  
41 84 90.3 17 4 US-09-346-847-17  
42 84 90.3 17 4 US-09-346-847-20  
43 84 90.3 17 4 US-09-346-847-22  
44 84 90.3 17 4 US-09-346-847-27  
45 84 90.3 17 4 US-09-648-400A-30

#### ALIGNMENTS

RESULT 1

US-08-928-958-7

; Sequence 7, Application US/08928958

; Patent No. 5877282

; GENERAL INFORMATION:

; APPLICANT: NADLER, STEVEN G.

; APPLICANT: CLEAVELAND, JEFFREY S.

; APPLICANT: BLAKE, JAMES

; APPLICANT: HAFFAR, OMAR K.

; TITLE OF INVENTION: PEPTIDE INHIBITORS OF NUCLEAR PROTEIN

; TITLE OF INVENTION: TRANSLLOCATION HAVING NUCLEAR LOCALIZATION SEQUENCES AND

; TITLE OF INVENTION: METHODS OF USE THEREOF

; NUMBER OF SEQUENCES: 24

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: ROBINS & ASSOCIATES

; STREET: 90 MIDDLEFIELD ROAD, SUITE 200

; CITY: MENLO PARK

; STATE: CA

; COUNTRY: USA

; ZIP: 94025

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/928,958

; FILING DATE: 12-SEP-1997

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 60/026978

; FILING DATE: 20-SEP-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: ROBINS, ROBERTA L.

; REGISTRATION NUMBER: 33,208

; REFERENCE/DOCKET NUMBER: 5998-0019

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (650) 325-7812

; TELEFAX: (650) 325-7823

; INFORMATION FOR SEQ ID NO: 7:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 16 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

US-08-928-958-7

Query Match 90.3%; Score 84; DB 2; Length 16;  
Best Local Similarity 93.8%; Pred. No. 3.5e-06;

Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RQIKWFOQRNRKWKFK 16  
Db 1 RQIKWFOQRNRKWKKK 16

RESULT 2

US-08-810-540-3  
; Sequence 3, Application US/08810540  
; Patent No. 5929042  
; GENERAL INFORMATION:  
; APPLICANT: Troy, Carol M.  
; APPLICANT: Shelanski, Michael L.  
; TITLE OF INVENTION: ANTISENSE COMPOUNDS WHICH PREVENT CELL  
; TITLE OF INVENTION: DEATH AND USES THEREOF  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Cooper & Dunham, LLP  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/810,540  
; FILING DATE: 03-MAR-1997  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White Esq., John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 0575/51247  
; TELEPHONE: 212-278-0400  
; TELEFAX: 212-391-0526  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 16 amino acids  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-810-540-3

Query Match 90.3%; Score 84; DB 2; Length 16;  
Best Local Similarity 93.8%; Pred. No. 3.5e-06;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RQIKWFOQRNRKWKFK 16  
Db 1 RQIKWFOQRNRKWKKK 16

RESULT 3

US-08-810-540-6  
; Sequence 6, Application US/08810540  
; Patent No. 5929042  
; GENERAL INFORMATION:  
; APPLICANT: Troy, Carol M.  
; APPLICANT: Shelanski, Michael L.  
; TITLE OF INVENTION: ANTISENSE COMPOUNDS WHICH PREVENT CELL  
; TITLE OF INVENTION: DEATH AND USES THEREOF  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Cooper & Dunham, LLP  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/072,429  
; FILING DATE: 04-MAY-1998  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Klein, Christopher A.  
; REGISTRATION NUMBER: 34,363  
; REFERENCE/DOCKET NUMBER: ONO141b  
; TELEPHONE: (609) 252-3714  
; TELEFAX: (609) 252-4526  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 16 amino acids  
; TYPE: amino acid

COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/810,540  
FILING DATE: 03-MAR-1997  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: White Esq., John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 0575/51247  
TELEPHONE: 212-278-0400  
TELEFAX: 212-391-0526  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 16 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-810-540-6

Query Match 90.3%; Score 84; DB 2; Length 16;  
Best Local Similarity 93.8%; Pred. No. 3.5e-06;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RQIKWFOQRNRKWKFK 16  
Db 1 RQIKWFOQRNRKWKKK 16

RESULT 4

US-09-072-429-7  
; Sequence 7, Application US/09072429  
; Patent No. 5962415  
; GENERAL INFORMATION:  
; APPLICANT: Nadler, Steven G.  
; TITLE OF INVENTION: COMPOSITIONS COMPRISING A PEPTIDE  
; TITLE OF INVENTION: INHIBITOR OF NUCLEAR PROTEIN TRANSLOCATION AND AN  
; TITLE OF INVENTION: IMMUNOSUPPRESSANT AND METHODS OF USE THEREOF  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Bristol-Myers Squibb Company  
; STREET: P.O. Box 4000  
; CITY: Princeton  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 08543-4000  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/072,429  
; FILING DATE: 04-MAY-1998  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Klein, Christopher A.  
; REGISTRATION NUMBER: 34,363  
; REFERENCE/DOCKET NUMBER: ONO141b  
; TELEPHONE: (609) 252-3714  
; TELEFAX: (609) 252-4526  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 16 amino acids  
; TYPE: amino acid



```
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-072-429-7

Query Match          90.3%; Score 84; DB 2; Length 16;
Best Local Similarity 93.8%; Pred. No. 3.5e-06;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RQIKWQNRRMKWFK 16
   |||||
Db 1 RQIKWQNRRMKWKK 16

RESULT 5
US-08-964-302A-6
; Sequence 6, Application US/08964302A
; Patent No. 6015787
; GENERAL INFORMATION:
; APPLICANT: Potter, David A.
; APPLICANT: Skolnik, Paul R.
; TITLE OF INVENTION: CELL-PERMEABLE PROTEIN INHIBITORS OF CALPAIN
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: Fast-SEQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/964.302A
; FILING DATE: 04-NOV-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Meiklejohn, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 00398/126001
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-964-302A-6

Query Match          90.3%; Score 84; DB 3; Length 16;
Best Local Similarity 93.8%; Pred. No. 3.5e-06;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RQIKWQNRRMKWFK 16
   |||||
Db 1 RQIKWQNRRMKWKK 16

RESULT 6
US-09-116-294-4
; Sequence 4, Application US/09116294
; Patent No. 6025140
; GENERAL INFORMATION:
; APPLICANT: Langel, Ulo
; APPLICANT: Bartfai, Tamas
; APPLICANT: Pooga, Margus
; APPLICANT: Valkna, Andres
; APPLICANT: Saar, Kulliki

; APPLICANT: Hallbrink, Mattias
; TITLE OF INVENTION: Conjugated Constructs of Peptides and
; TITLE OF INVENTION: Nucleic Acid Analogs, and Their Transport Across Membranes
; FILE REFERENCE: 4394
; CURRENT APPLICATION NUMBER: US/09/116.294
; CURRENT FILING DATE: 1998-07-16
; EARLIER APPLICATION NUMBER: 60/052.678
; EARLIER FILING DATE: 1997-07-24
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 16
; TYPE: PPT
; ORGANISM: drosophila
US-09-116-294-4

Query Match          90.3%; Score 84; DB 3; Length 16;
Best Local Similarity 93.8%; Pred. No. 3.5e-06;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RQIKWQNRRMKWFK 16
   |||||
Db 1 RQIKWQNRRMKWKK 16

RESULT 7
US-08-964-614A-4
; Sequence 4, Application US/08964614A
; Patent No. 6057104
; GENERAL INFORMATION:
; APPLICANT: Hasty, Paul
; TITLE OF INVENTION: DISRUPTION OF THE MAMMALIAN
; TITLE OF INVENTION: Rad51 PROTEIN AND DISRUPTION OF PROTEINS THAT ASSOCIATE
; TITLE OF INVENTION: WITH MAMMALIAN Rad51 FOR HINDERING CELL PROLIFERATION
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds, LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: Fast-SEQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/964.614A
; FILING DATE: 05-NOV-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/758,280
; FILING DATE: 05-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 8535-0019-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-493-4935
; TELEFAX: 650-493-5556
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-964-614A-4

Query Match          90.3%; Score 84; DB 3; Length 16;
Best Local Similarity 93.8%; Pred. No. 3.5e-06;
```

Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RQIKWQNRRMKWFK 16  
| | | | | | | | | | | | | | | | | |  
Db 1 RQIKWQNRRMKWKK 16

RESULT 8  
US-08-849-486-1  
; Sequence 1, Application US/08849486  
; Patent No. 6080724  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: PEPTIDES WHICH CAN BE USED AS VECTORS  
; TITLE OF INVENTION: FOR THE INTRACELLULAR ADDRESSING OF ACTIVE MOLECULES  
; NUMBER OF SEQUENCES: 10  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION NUMBER: US/08/849,486  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: FR 95 11714  
; FILING DATE: 05-OCT-1995  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 16 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-849-486-1

Query Match 90.3%; Score 84; DB 3; Length 16;  
Best Local Similarity 93.8%; Pred. No. 3.5e-06;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RQIKWQNRRMKWFK 16  
| | | | | | | | | | | | | | | | | |  
Db 1 RQIKWQNRRMKWKK 16

RESULT 9  
US-08-849-486-4  
; Sequence 4, Application US/08849486  
; Patent No. 6080724  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: PEPTIDES WHICH CAN BE USED AS VECTORS  
; TITLE OF INVENTION: FOR THE INTRACELLULAR ADDRESSING OF ACTIVE MOLECULES  
; NUMBER OF SEQUENCES: 10  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION NUMBER: US/08/849,486  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: FR 95 11714  
; FILING DATE: 05-OCT-1995  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 16 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear

; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: Peptide  
; LOCATION: 1..16  
; OTHER INFORMATION: /product= "amino acids of the D series"  
US-08-849-486-4

Query Match 90.3%; Score 84; DB 3; Length 16;  
Best Local Similarity 93.8%; Pred. No. 3.5e-06;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RQIKWQNRRMKWFK 16  
| | | | | | | | | | | | | | | | | |  
Db 1 RQIKWQNRRMKWKK 16

RESULT 10  
US-09-208-966-54  
; Sequence 54, Application US/09208966  
; Patent No. 6221355  
; GENERAL INFORMATION:  
; APPLICANT: Dowdy, Steven F.  
; TITLE OF INVENTION: ANTI-PATHOGEN SYSTEM AND METHODS OF USE THEREOF  
; FILE REFERENCE: 48881/1742  
; CURRENT APPLICATION NUMBER: US/09/208,966  
; CURRENT FILING DATE: 1998-12-10  
; EARLIER APPLICATION NUMBER: 60/082,402  
; EARLIER FILING DATE: 1998-04-20  
; EARLIER APPLICATION NUMBER: 60/069,012  
; EARLIER FILING DATE: 1997-12-10  
; NUMBER OF SEQ ID NOS: 57  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 54  
; LENGTH: 16  
; TYPE: PPT  
; ORGANISM: human  
US-09-208-966-54

Query Match 90.3%; Score 84; DB 3; Length 16;  
Best Local Similarity 93.8%; Pred. No. 3.5e-06;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RQIKWQNRRMKWFK 16  
| | | | | | | | | | | | | | | | | |  
Db 1 RQIKWQNRRMKWKK 16

RESULT 11  
US-09-308-935-8  
; Sequence 8, Application US/09308935  
; Patent No. 6268334  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas B  
; APPLICANT: Bandara, Laxantha R  
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
; FILE REFERENCE: 620-67  
; CURRENT APPLICATION NUMBER: US/09/308,935  
; CURRENT FILING DATE: 1999-05-27  
; EARLIER APPLICATION NUMBER: PCT/GB97/03506  
; EARLIER FILING DATE: 1997-12-22  
; EARLIER APPLICATION NUMBER: GB 9626589.7  
; EARLIER FILING DATE: 1996-12-20  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 8  
; LENGTH: 16  
; TYPE: PPT  
; ORGANISM: Drosophila melanogaster  
US-09-308-935-8

Query Match 90.3%; Score 84; DB 3; Length 16;  
Best Local Similarity 93.8%; Pred. No. 3.5e-06;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy 1 RQIKWQNRRMKWFK 16
Db 1 RQIKWQNRRMKWKK 16

RESULT 12
US-09-441-416A-6
; Sequence 6, Application US/09441416A
; Patent No. 6294518
; GENERAL INFORMATION:
; APPLICANT: Potter, David A.
; APPLICANT: Skolnik, Paul R.
; TITLE OF INVENTION: CELL-PERMEABLE PROTEIN INHIBITORS OF
; TITLE OF INVENTION: CALPAIN
; FILE REFERENCE: 00398-140001
; CURRENT APPLICATION NUMBER: US/09/441,416A
; CURRENT FILING DATE: 1999-11-16
; PRIOR APPLICATION NUMBER: US 08/964,302
; PRIOR FILING DATE: 1997-11-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-441-416A-6

Query Match 90.3%; Score 84; DB 3; Length 16;
Best Local Similarity 93.8%; Pred. No. 3.5e-06;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RQIKWQNRRMKWFK 16
Db 1 RQIKWQNRRMKWKK 16

RESULT 13
US-09-296-089-33
; Sequence 33, Application US/09296089
; Patent No. 6303576
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Byers, Stephen
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING
; TITLE OF INVENTION: BETA-CATENIN MEDIATED GENE EXPRESSION
; FILE REFERENCE: 100086.411
; CURRENT APPLICATION NUMBER: US/09/296,089
; CURRENT FILING DATE: 1999-04-21
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 33
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-296-089-33

Query Match 90.3%; Score 84; DB 3; Length 16;
Best Local Similarity 93.8%; Pred. No. 3.5e-06;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RQIKWQNRRMKWFK 16
Db 1 RQIKWQNRRMKWKK 16

RESULT 14
US-09-419-826-35
; Sequence 35, Application US/09419826
; Patent No. 6306832
; GENERAL INFORMATION:
; APPLICANT:
```

```
; TITLE OF INVENTION: PEPTIDE ANTIESTROGEN COMPOSITIONS AND METHODS
; TITLE OF INVENTION: FOR TREATING BREAST CANCER
; NUMBER OF SEQUENCES: 39
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/419,826
; FILING DATE: 14-OCT-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US98/07711
; FILING DATE: 14-APR-1998
; APPLICATION NUMBER: US 60/043,545
; FILING DATE: 14-APR-1997
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-09-419-826-35
```

Query Match 90.3%; Score 84; DB 3; Length 16;  
Best Local Similarity 93.8%; Pred. No. 3.5e-06;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy 1 RQIKWQNRRMKWFK 16
Db 1 RQIKWQNRRMKWKK 16
```

```
RESULT 15
US-09-302-305C-10
; Sequence 10, Application US/09302305C
; Patent No. 6350572
; GENERAL INFORMATION:
; APPLICANT: Bernards, Rene
; APPLICANT: Zwijssen, Renate
; TITLE OF INVENTION: Interaction Between Cyclin D1 and Steroid Receptor
; TITLE OF INVENTION: Co-Activators and Uses Thereof in Assays
; FILE REFERENCE: 4238/80713
; CURRENT APPLICATION NUMBER: US/09/302,305C
; CURRENT FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: PCT/GB99/00440
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 10
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(16)
; OTHER INFORMATION: Translocation peptide derived from antennapedia
; OTHER INFORMATION: homeodomain protein
US-09-302-305C-10
```

Query Match 90.3%; Score 84; DB 3; Length 16;  
Best Local Similarity 93.8%; Pred. No. 3.5e-06;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy 1 RQIKWQNRRMKWFK 16
Db 1 RQIKWQNRRMKWKK 16
```

Search completed: March 31, 2005, 00:32:41  
Job time : 5.2 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2005, 00:25:06 ; Search time 7.69231 Seconds  
(without alignments)  
688.690 Million cell updates/sec

Title: US-10-049-822A-13

Perfect score: 93  
Sequence: 1 RQIKWQNRRMKWFK 16

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1407402 seqs, 331100923 residues

Total number of hits satisfying chosen parameters: 1407402

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*

2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*

3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*

4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*

5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*

6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*

7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*

8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*

9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*

10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*

11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*

12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*

13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*

14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*

15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*

16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*

17: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*

18: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*

19: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*

20: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	85	91.4	17	9	US-09-854-204-34
2	84	90.3	16	8	US-08-610-220A-9
3	84	90.3	16	9	US-09-748-063-3
4	84	90.3	16	9	US-09-214-371-43
5	84	90.3	16	9	US-09-779-791A-3
6	84	90.3	16	9	US-09-780-070-38
7	84	90.3	16	9	US-09-150-623-9
8	84	90.3	16	9	US-09-731-023A-10
9	84	90.3	16	9	US-09-854-204-1
10	84	90.3	16	9	US-09-900-147-8
11	84	90.3	16	9	US-09-792-480-29
12	84	90.3	16	9	US-09-785-802A-2
13	84	90.3	16	9	US-09-785-802A-5

14	84	90.3	16	9	US-09-902-432-32
15	84	90.3	16	9	US-09-953-031A-10
16	84	90.3	16	9	US-09-981-286A-3
17	84	90.3	16	10	US-09-962-967A-6
18	84	90.3	16	10	US-09-912-414-6
19	84	90.3	16	10	US-09-775-052-54
20	84	90.3	16	10	US-09-295-189-4
21	84	90.3	16	10	US-09-965-876A-1
22	84	90.3	16	10	US-09-933-780C-2
23	84	90.3	16	12	US-09-779-791A-3
24	84	90.3	16	13	US-10-024-935-12
25	84	90.3	16	13	US-10-007-363-3
26	84	90.3	16	14	US-10-083-960-29
27	84	90.3	16	14	US-10-071-512A-2
28	84	90.3	16	14	US-10-239-804-3
29	84	90.3	16	14	US-10-077-555-3
30	84	90.3	16	14	US-10-209-421-29
31	84	90.3	16	14	US-10-229-915-2
32	84	90.3	16	14	US-10-185-084-3
33	84	90.3	16	14	US-10-252-012-5
34	84	90.3	16	14	US-10-075-869-19
35	84	90.3	16	14	US-10-013-815-19
36	84	90.3	16	14	US-10-136-738-10
37	84	90.3	16	14	US-10-210-660-1
38	84	90.3	16	14	US-10-210-660-25
39	84	90.3	16	14	US-10-156-570A-21
40	84	90.3	16	14	US-10-201-394A-14
41	84	90.3	16	14	US-10-017-672-11
42	84	90.3	16	14	US-10-201-389A-14
43	84	90.3	16	14	US-10-161-051-1
44	84	90.3	16	14	US-10-358-365-10
45	84	90.3	16	14	US-10-061-607A-2

ALIGNMENTS

RESULT 1

US-09-854-204-34

; Sequence 34, Application US/09854204

; Patent No. US20020098236A1

; GENERAL INFORMATION:

; APPLICANT: Fischer, Peter Martin

; APPLICANT: Zhelev, Nikolai

; TITLE OF INVENTION: Transport Vectors

; FILE REFERENCE: CCI-010

; CURRENT APPLICATION NUMBER: US/09/854,204

; CURRENT FILING DATE: 2001-05-11

; PRIOR APPLICATION NUMBER: 09/438,460

; PRIOR FILING DATE: 1999-11-12

; PRIOR APPLICATION NUMBER: GB 9825000.4

; PRIOR FILING DATE: 1998-11-13

; PRIOR APPLICATION NUMBER: GB 9825001.2

; PRIOR FILING DATE: 1998-11-13

; PRIOR APPLICATION NUMBER: GB 9902525.6

; PRIOR FILING DATE: 1999-02-04

; PRIOR APPLICATION NUMBER: GB 9902522.3

; PRIOR FILING DATE: 1999-02-04

; PRIOR APPLICATION NUMBER: GB 9914578.1

; PRIOR FILING DATE: 1999-06-22

; PRIOR APPLICATION NUMBER: PCT/GB99/03750

; PRIOR FILING DATE: 1999-11-11

; NUMBER OF SEQ ID NOS: 66

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 34

; LENGTH: 17

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; NAME/KEY: MOD\_RES

; LOCATION: (1)

; OTHER INFORMATION: bala

; NAME/KEY: MOD\_RES

; LOCATION: (17)  
; OTHER INFORMATION: AMIDATION  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: sequence  
US-09-854-204-34

Query Match 91.4%; Score 85; DB 9; Length 17;  
Best Local Similarity 93.8%; Pred. No. 2.6e-05;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RQIKWQNRRMKWFK 16  
| | | | | | | | | | | | | | | | | |  
Db 2 RQIKWQNRRMKWAK 17

RESULT 2  
US-08-610-220A-9  
; Sequence 9, Application US/08610220A  
; Publication No. US2003009638A1  
; GENERAL INFORMATION:  
; APPLICANT: Troy, Carol M.  
; TITLE OF INVENTION: COMPOUNDS WHICH PREVENT NEURONAL CELL  
; TITLE OF INVENTION: DEATH AND USES THEREOF  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper & Dunham LLP  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/610,220A  
; FILING DATE: MAR-04-1996  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 48332/JPW/JML  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-278-0400  
; TELEFAX: 212-391-0525  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 16 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-610-220A-9

Query Match 90.3%; Score 84; DB 8; Length 16;  
Best Local Similarity 93.8%; Pred. No. 3.4e-05;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RQIKWQNRRMKWFK 16  
| | | | | | | | | | | | | | | | | |  
Db 1 RQIKWQNRRMKWKK 16

RESULT 3  
US-09-748-063-3  
; Sequence 3, Application US/09748063  
; Publication No. US20010008758A1  
; GENERAL INFORMATION:  
; APPLICANT: McHale, Anthony P.  
; APPLICANT: Craig, Roger  
; APPLICANT: Haro, Anna Maria Rollan

; TITLE OF INVENTION: Delivery of an Agent  
; FILE REFERENCE: 11067/1060  
; CURRENT APPLICATION NUMBER: US/09/748,063  
; CURRENT FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: PCT/GB00/02848  
; PRIOR FILING DATE: 2000-07-24  
; PRIOR APPLICATION NUMBER: US 60/146,556  
; PRIOR FILING DATE: 2000-07-30  
; PRIOR APPLICATION NUMBER: GB 9917416.1  
; PRIOR FILING DATE: 1999-07-23  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 3  
; LENGTH: 16  
; TYPE: PRT  
; ORGANISM: Drosophila sp.  
US-09-748-063-3

Query Match 90.3%; Score 84; DB 9; Length 16;  
Best Local Similarity 93.8%; Pred. No. 3.4e-05;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RQIKWQNRRMKWFK 16  
| | | | | | | | | | | | | | | | | |  
Db 1 RQIKWQNRRMKWKK 16

RESULT 4  
US-09-214-371-43  
; Sequence 43, Application US/09214371B  
; Patent No. US20010018511A1  
; GENERAL INFORMATION:  
; APPLICANT: Lane, David  
; APPLICANT: Bottger, Volker  
; APPLICANT: Bottger, Angelica  
; APPLICANT: Picklesley, Stephen  
; APPLICANT: Chene, Patrick  
; APPLICANT: Hochkeppel, Heinz-Kurt  
; APPLICANT: Garcia-Echeverria, Carlos  
; APPLICANT: Furet, Pascal  
; TITLE OF INVENTION: Inhibitors of the Interaction of P53 and MDM2  
; FILE REFERENCE: 4-20937/A/PCT  
; CURRENT APPLICATION NUMBER: US/09/214,371B  
; CURRENT FILING DATE: 1999-03-26  
; PRIOR APPLICATION NUMBER: PCT/EP97/03549  
; PRIOR FILING DATE: 1997-07-04  
; NUMBER OF SEQ ID NOS: 83  
; SOFTWARE: Patent in Ver. 2.0  
; SEQ ID NO 43  
; LENGTH: 16  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: peptide  
US-09-214-371-43

Query Match 90.3%; Score 84; DB 9; Length 16;  
Best Local Similarity 93.8%; Pred. No. 3.4e-05;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RQIKWQNRRMKWFK 16  
| | | | | | | | | | | | | | | | | |  
Db 1 RQIKWQNRRMKWKK 16

RESULT 5  
US-09-779-791A-3  
; Sequence 3, Application US/09779791A  
; Publication No. US2001004417A1  
; GENERAL INFORMATION:  
; APPLICANT: Mirus Corporation  
; APPLICANT: Wolff, Jon A  
; APPLICANT: Monahan, Sean D

APPLICANT: Budker, Vladimir G  
APPLICANT: Slatum, Paul M  
APPLICANT: Rozema, David B  
TITLE OF INVENTION: A Compound Containing a Labile Disulfide Bond  
FILE REFERENCE: Mirus.006.03  
CURRENT APPLICATION NUMBER: US/09/779,791A  
CURRENT FILING DATE: 2001-02-08  
PRIOR APPLICATION NUMBER: 09/312,351  
PRIOR FILING DATE: 1999-05-14  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 3  
LENGTH: 16  
TYPE: PRT  
ORGANISM: Drosophila melanogaster  
US-09-779-791A-3

Query Match 90.3%; Score 84; DB 9; Length 16;  
Best Local Similarity 93.8%; Pred. No. 3.4e-05;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RQIKWQNRRMKWFK 16  
| | | | | | | | | | | | | | | |  
Db 1 RQIKWQNRRMKWKK 16

RESULT 6  
US-09-780-070-38  
Sequence 38, Application US/09780070  
Patent No. US2002009752A1  
GENERAL INFORMATION:  
APPLICANT: Burke, James  
APPLICANT: Strittmatter, Warren  
APPLICANT: Nagai, Yoshitaka  
TITLE OF INVENTION: COMPOUNDS THAT SELECTIVELY BIND TO EXPANDED POLYGLUTAMINE REPEAT  
TITLE OF INVENTION: AND METHODS OF USE THEREOF  
FILE REFERENCE: 5405.242  
CURRENT APPLICATION NUMBER: US/09/780,070  
CURRENT FILING DATE: 2001-02-09  
PRIOR APPLICATION NUMBER: 60/189,781  
PRIOR FILING DATE: 2000-03-16  
NUMBER OF SEQ ID NOS: 40  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 38  
LENGTH: 16  
TYPE: PRT  
ORGANISM: Drosophila melanogaster  
US-09-780-070-38

Query Match 90.3%; Score 84; DB 9; Length 16;  
Best Local Similarity 93.8%; Pred. No. 3.4e-05;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RQIKWQNRRMKWFK 16  
| | | | | | | | | | | | | | | |  
Db 1 RQIKWQNRRMKWKK 16

RESULT 7  
US-09-150-623-9  
Sequence 9, Application US/09150623  
Patent No. US2002004931A1  
GENERAL INFORMATION:  
APPLICANT: Troy, Carol M.  
TITLE OF INVENTION: COMPOUNDS WHICH PREVENT NEURONAL CELL  
DEATH AND USES THEREOF  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooper & Dunham LLP  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.

ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/150,623  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/610,220  
FILING DATE: MAR-04-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 48332/JPW/JML  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-278-0400  
TELEFAX: 212-391-0525  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 16 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-150-623-9

Query Match 90.3%; Score 84; DB 9; Length 16;  
Best Local Similarity 93.8%; Pred. No. 3.4e-05;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RQIKWQNRRMKWFK 16  
| | | | | | | | | | | | | | | |  
Db 1 RQIKWQNRRMKWKK 16

RESULT 8  
US-09-731-023A-10  
Sequence 10, Application US/09731023A  
Patent No. US2002007283A1  
GENERAL INFORMATION:  
APPLICANT: Sessa, William  
TITLE OF INVENTION: Caveolin Peptides and Their Use as Therapeutics  
FILE REFERENCE: 44574-5076-US  
CURRENT APPLICATION NUMBER: US/09/731,023A  
CURRENT FILING DATE: 2000-12-07  
PRIOR APPLICATION NUMBER: US 60/231,327  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 10  
LENGTH: 16  
TYPE: PRT  
ORGANISM: Drosophila melanogaster  
FEATURE:  
NAME/KEY: DOMAIN  
LOCATION: (1)...(16)  
OTHER INFORMATION: Homeodomain, internalization sequence  
US-09-731-023A-10

Query Match 90.3%; Score 84; DB 9; Length 16;  
Best Local Similarity 93.8%; Pred. No. 3.4e-05;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RQIKWQNRRMKWFK 16  
| | | | | | | | | | | | | | | |  
Db 1 RQIKWQNRRMKWKK 16

RESULT 9  
US-09-854-204-1

US-09-792-480-29  
; Sequence 1, Application US/09854204  
; Patent No. US20020098236A1  
; GENERAL INFORMATION:  
; APPLICANT: Fischer, Peter Martin  
; APPLICANT: Zhelev, Nikolai  
; TITLE OF INVENTION: Transport Vectors  
; FILE REFERENCE: CCI-010  
; CURRENT APPLICATION NUMBER: US/09/854,204  
; CURRENT FILING DATE: 2001-05-11  
; PRIOR APPLICATION NUMBER: 09/438,460  
; PRIOR FILING DATE: 1999-11-12  
; PRIOR APPLICATION NUMBER: GB 9825000.4  
; PRIOR FILING DATE: 1998-11-13  
; PRIOR APPLICATION NUMBER: GB 9825001.2  
; PRIOR FILING DATE: 1998-11-13  
; PRIOR APPLICATION NUMBER: GB 9902525.6  
; PRIOR FILING DATE: 1999-02-04  
; PRIOR APPLICATION NUMBER: GB 9902522.3  
; PRIOR FILING DATE: 1999-02-04  
; PRIOR APPLICATION NUMBER: GB 9914578.1  
; PRIOR FILING DATE: 1999-06-22  
; PRIOR APPLICATION NUMBER: PCT/GB99/03750  
; PRIOR FILING DATE: 1999-11-11  
; NUMBER OF SEQ ID NOS: 66  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 16  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
US-09-854-204-1

Query Match 90.3%; Score 84; DB 9; Length 16;  
Best Local Similarity 93.8%; Pred. No. 3.4e-05;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 1 RQIKWQNRRMKWFK 16  
| | | | | | | | | | | | | | | |  
Db 1 RQIKWQNRRMKWKK 16

RESULT 10  
US-09-900-147-8  
; Sequence 8, Application US/09900147  
; Patent No. US20020103121A1  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas B  
; APPLICANT: Bandara, Lasantha R  
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
; FILE REFERENCE: 620-67  
; CURRENT APPLICATION NUMBER: US/09/900,147  
; CURRENT FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/308,935  
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-05-27  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: GB 9626589.7  
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-12-20  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 8  
; LENGTH: 16  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
US-09-900-147-8

Query Match 90.3%; Score 84; DB 9; Length 16;  
Best Local Similarity 93.8%; Pred. No. 3.4e-05;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 1 RQIKWQNRRMKWFK 16  
| | | | | | | | | | | | | | | |  
Db 1 RQIKWQNRRMKWKK 16

RESULT 11

US-09-792-480-29  
; Sequence 29, Application US/09792480  
; Patent No. US20020127198A1  
; GENERAL INFORMATION:  
; APPLICANT: Rothbard, Jonathan B.  
; APPLICANT: Wender, Paul A.  
; APPLICANT: McGrane, P. Leo  
; APPLICANT: Sista, Lalitha V.S.  
; APPLICANT: Kirschberg, Thorsten A.  
; APPLICANT: CellGate, Inc.  
; TITLE OF INVENTION: Compositions and Methods for Enhancing Drug Delivery  
; FILE REFERENCE: 019801-000230US  
; CURRENT APPLICATION NUMBER: US/09/792,480  
; CURRENT FILING DATE: 2001-02-23  
; PRIOR APPLICATION NUMBER: US 09/648,400  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/150,510  
; PRIOR FILING DATE: 1999-08-24  
; NUMBER OF SEQ ID NOS: 57  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 29  
; LENGTH: 16  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:Antennapedia  
; OTHER INFORMATION: homeodomain, Antennapedia-43-58  
US-09-792-480-29

Query Match 90.3%; Score 84; DB 9; Length 16;  
Best Local Similarity 93.8%; Pred. No. 3.4e-05;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 1 RQIKWQNRRMKWFK 16  
| | | | | | | | | | | | | | | |  
Db 1 RQIKWQNRRMKWKK 16

RESULT 12  
US-09-785-802A-2  
; Sequence 2, Application US/09785802A  
; Patent No. US20020151004A1  
; GENERAL INFORMATION:  
; APPLICANT: Craig, Roger  
; TITLE OF INVENTION: DELIVERY VEHICLES AND METHODS FOR USING THE SAME  
; FILE REFERENCE: 11067/2035  
; CURRENT APPLICATION NUMBER: US/09/785,802A  
; CURRENT FILING DATE: 2001-02-16  
; PRIOR APPLICATION NUMBER: US 09/748,06  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: US 09/748,789  
; PRIOR FILING DATE: 2000-12-22  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 16  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Membrane translocation sequence from Penetratin  
US-09-785-802A-2

Query Match 90.3%; Score 84; DB 9; Length 16;  
Best Local Similarity 93.8%; Pred. No. 3.4e-05;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 1 RQIKWQNRRMKWFK 16  
| | | | | | | | | | | | | | | |  
Db 1 RQIKWQNRRMKWKK 16

RESULT 13



```
US-09-785-802A-5
; Sequence 5, Application US/09785802A
; Patent No. US20020151004A1
; GENERAL INFORMATION:
; APPLICANT: Craig, Roger
; TITLE OF INVENTION: DELIVERY VEHICLES AND METHODS FOR USING THE SAME
; FILE REFERENCE: 11067/2035
; CURRENT APPLICATION NUMBER: US/09/785.802A
; CURRENT FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 09/748,06
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/748,789
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 5
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-785-802A-5
Query Match          90.3%; Score 84; DB 9; Length 16;
Best Local Similarity 93.8%; Pred. No. 3.4e-05;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RQIKWQNRRMKWFK 16
   |||||
Db 1 RQIKWQNRRMKWKK 16

RESULT 14
US-09-902-432-32
; Sequence 32, Application US/09902432
; Patent No. US2002016002A1
; GENERAL INFORMATION:
; APPLICANT: Irwin H. Gelman
; APPLICANT: Susan G. Jaken
; TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
; FILE REFERENCE: A30558-A-FWC-A 070156.0597
; CURRENT APPLICATION NUMBER: US/09/902,432
; CURRENT FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 08/978,277
; PRIOR FILING DATE: 1997-11-25
; PRIOR APPLICATION NUMBER: 08/665,401
; PRIOR FILING DATE: 1996-06-18
; PRIOR APPLICATION NUMBER: 08/635,121
; PRIOR FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Penetratin peptide
US-09-902-432-32
Query Match          90.3%; Score 84; DB 9; Length 16;
Best Local Similarity 93.8%; Pred. No. 3.4e-05;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RQIKWQNRRMKWFK 16
   |||||
Db 1 RQIKWQNRRMKWKK 16

RESULT 15
US-09-953-031A-10
; Sequence 10, Application US/09953031A
; Patent No. US2002017717A1
; GENERAL INFORMATION:
; APPLICANT: Bernards, Rene
; APPLICANT: Zwijssen, Renate
```

```
; TITLE OF INVENTION: Interaction Between Cyclin D1 and Steroid Receptor
; TITLE OF INVENTION: Co-Activators and Uses Thereof in Assays
; FILE REFERENCE: 4238/80713
; CURRENT APPLICATION NUMBER: US/09/953.031A
; CURRENT FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: US 09/302,305
; PRIOR FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 10
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(16)
; OTHER INFORMATION: Translocation peptide derived from antennapedia
; OTHER INFORMATION: homeodomain protein
US-09-953-031A-10
Query Match          90.3%; Score 84; DB 9; Length 16;
Best Local Similarity 93.8%; Pred. No. 3.4e-05;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RQIKWQNRRMKWFK 16
   |||||
Db 1 RQIKWQNRRMKWKK 16

Search completed: March 31, 2005, 00:34:58
Job time : 9.69231 secs
```

100

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2005, 00:22:11 ; Search time 2.09231 Seconds  
(without alignments)  
735.775 Million cell updates/sec

Title: US-10-049-822a-13  
Perfect score: 93  
Sequence: 1 RQIKWIFQNRMRKWK 16

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR\_79:\*  
1: Pirl.\*  
2: Pirl2.\*  
3: Pirl3.\*  
4: Pirl4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	84	90.3	33	2 S57235	antennapedia prote
2	84	90.3	42	2 I65241	homeotic protein H
3	84	90.3	45	2 PC1216	homeotic protein D
4	84	90.3	48	2 I51439	homeobox protein -
5	84	90.3	50	2 S10907	homeotic protein (
6	84	90.3	66	2 S15538	homeotic protein H
7	84	90.3	66	2 S15536	homeotic protein H
8	84	90.3	71	2 JC1161	homeotic protein 3
9	84	90.3	71	2 A60084	homeotic protein H
10	84	90.3	74	2 D34510	homeotic protein H
11	84	90.3	75	2 S88852	homeotic protein S
12	84	90.3	75	2 I51341	homeo box protein
13	84	90.3	76	2 C43559	homeo box protein R
14	84	90.3	78	2 I51342	homeo box protein H
15	84	90.3	81	2 B29585	homeotic protein H
16	84	90.3	81	2 S47605	homeotic protein H
17	84	90.3	82	2 S08302	homeotic protein H
18	84	90.3	83	2 S50066	homeotic protein H
19	84	90.3	83	2 S47603	homeotic protein H
20	84	90.3	86	2 A34510	homeotic protein H
21	84	90.3	86	2 S08303	homeotic protein H
22	84	90.3	86	2 JT0489	homeotic protein Z
23	84	90.3	87	2 S00589	homeotic protein M
24	84	90.3	88	2 A03317	homeotic protein M
25	84	90.3	96	2 S08639	homeotic protein z
26	84	90.3	96	2 A05266	homeotic protein H
27	84	90.3	97	2 C27176	homeotic protein H
28	84	90.3	97	2 A24779	homeotic protein m
29	84	90.3	103	2 A32167	homeotic protein H

30	84	90.3	105	2 S47602	homeotic protein H
31	84	90.3	105	2 A27471	homeotic protein R
32	84	90.3	106	2 S36448	homeotic protein s
33	84	90.3	107	2 B61045	homeotic protein T
34	84	90.3	113	2 T10775	homeobox protein -
35	84	90.3	118	2 A24777	homeotic protein H
36	84	90.3	118	2 JT0273	homeotic protein H
37	84	90.3	118	2 B24777	homeotic protein M
38	84	90.3	119	2 A03314	homeotic protein m
39	84	90.3	138	2 S20087	homeotic protein b
40	84	90.3	148	2 PC4071	homeobox A5 protei
41	84	90.3	153	1 WJHU3C	homeotic protein H
42	84	90.3	153	1 WJMSX6	homeotic protein H
43	84	90.3	158	2 A27348	homeotic protein H
44	84	90.3	209	2 A43553	homeotic protein H
45	84	90.3	217	1 WJHU2C	homeotic protein H

ALIGNMENTS

RESULT 1

S57235  
antennapedia protein (clone pl105) - fruit fly (Drosophila pseudoobscura) (fragment)  
C:Species: Drosophila pseudoobscura  
C:Date: 10-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 16-Aug-2004  
C:Accession: S57235  
R:Randazzo, F.M.; Seeger, M.A.; Huss, C.A.; Sweeney, M.A.; Cecil, J.K.; Kaufman, T.C.  
Genetics 133, 319-330, 1993  
A:Title: Structural changes in the antennapedia complex of Drosophila pseudoobscura.  
A:Reference number: S57224  
A:Accession: S57235  
A:Molecule type: DNA  
A:Residues: 1-33 <RAN>  
A:Cross-references: EMBL:X77711  
C:Genetics:  
A:Gene: FlyBase:Antp  
A:Cross-references: FlyBase:FBgn0012693  
C:Superfamily: homeobox homology  
C:Keywords: DNA binding; homeobox; nucleus; transcription regulation  
F:1-22/Domain: homeobox homology (fragment) <HOX>

Query Match 90.3%; Score 84; DB 2; Length 33;  
Best Local Similarity 93.8%; Pred. No. 3.1e-06;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RQIKWIFQNRMRKWK 16  
| | | | | | | | | | | | | | | |  
Db 7 RQIKWIFQNRMRKWK 22

RESULT 2

I65241  
homeotic protein Hox-A - rat (fragment)  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 27-Feb-1997 #sequence\_revision 27-Feb-1997 #text\_change 16-Aug-2004  
C:Accession: I65241  
R:Sakoyama, Y.; Mizuta, I.; Ogasawara, N.; Yoshikawa, H.  
Biochem. Genet. 32, 351-360, 1994  
A:Title: Cloning of rat homeobox genes.  
A:Reference number: 152340; MUID:95217128; PMID:7702549  
A:Accession: I65241  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-42 <RES>  
A:Cross-references: GB:S76290; NID:g913077  
C:Genetics:  
A:Gene: Hox-A; Hox-1  
C:Superfamily: homeobox homology  
C:Keywords: DNA binding; homeobox; nucleus; transcription regulation  
F:1-40/Domain: homeobox homology (fragment) <HOX>

Query Match 90.3%; Score 84; DB 2; Length 42;

Best Local Similarity 93.8%; Pred. No. 3.9e-06;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RQIKWQNRRMKWFK 16  
| | | | | | | | | | | | | | | | | |  
Db 25 RQIKWQNRRMKWKK 40

RESULT 3  
PC1216  
homeotic protein Dtbx1 - planarian (*Dugesia tigrina*) (fragment)  
C:Species: *Dugesia tigrina*  
C>Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Aug-2004  
C:Accession: PC1216  
R:Oliver, G.; Vispo, M.; Mailhos, A.; Martinez, C.; Sosa-Pineda, B.; Flietitz, W.; Ehrlich  
Gene 121, 337-342, 1992  
A:Title: Homeoboxes in flatworms.  
A:Reference number: JC1386; MUID:93077050; PMID:1359988  
A:Accession: PC1216  
A:Molecule type: DNA  
A:Residues: 1-45 <OLI>  
A:Cross-references: UNIPROT:P91714; EMBL:X66822  
C:Superfamily: homeobox homology  
C:Keywords: DNA binding; homeobox; nucleus; transcription regulation  
F;1-45/Domain: homeobox homology (fragment) <HOX>

Query Match 90.3%; Score 84; DB 2; Length 45;  
Best Local Similarity 93.8%; Pred. No. 4.1e-06;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RQIKWQNRRMKWFK 16  
| | | | | | | | | | | | | | | | | |  
Db 30 RQIKWQNRRMKWKK 45

RESULT 4  
I51439  
homeobox protein - African clawed frog (fragment)  
C:Species: *Xenopus laevis* (African clawed frog)  
C>Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 16-Aug-2004  
C:Accession: I51439  
R:Leroy, P.; DeRobertis, E.M.  
Dev. Dyn. 194, 21-32, 1992  
A:Title: Effects of lithium chloride and retinoic acid on the expression of genes from t  
A:Reference number: I51439; MUID:93043517; PMID:1384809  
A:Accession: I51439  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-48 <LER>  
A:Cross-references: UNIPROT:P31256; GB:M91587; NID:9214257; PIDN:AAA49750.1; PID:9214258  
C:Genetics:  
A:Gene: Hox2.2  
C:Superfamily: homeobox homology  
C:Keywords: DNA binding; homeobox; nucleus; transcription regulation  
F;1-27/Domain: homeobox homology (fragment) <HOX>

Query Match 90.3%; Score 84; DB 2; Length 48;  
Best Local Similarity 93.8%; Pred. No. 4.4e-06;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RQIKWQNRRMKWFK 16  
| | | | | | | | | | | | | | | | | |  
Db 12 RQIKWQNRRMKWKK 27

RESULT 5  
S10907  
homeotic protein (clone 11-2-1) - human (fragment)  
C:Species: *Homo sapiens* (man)  
C>Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 16-Aug-2004  
C:Accession: S10907  
R:Gilman, J.G.  
submitted to the EMBL Data Library, April 1990

A:Reference number: S10907  
A:Accession: S10907  
A:Molecule type: mRNA  
A:Residues: 1-50 <GIL>  
A:Cross-references: UNIPROT:Q14543; EMBL:X52402; NID:932360; PIDN:CAA36647.1; PID:932361  
C:Superfamily: homeobox homology  
C:Keywords: DNA binding; homeobox; nucleus; transcription regulation  
F;1-40/Domain: homeobox homology (fragment) <HOX>

Query Match 90.3%; Score 84; DB 2; Length 50;  
Best Local Similarity 93.8%; Pred. No. 4.6e-06;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RQIKWQNRRMKWFK 16  
| | | | | | | | | | | | | | | | | |  
Db 25 RQIKWQNRRMKWKK 40

RESULT 6  
S15538  
homeotic protein Hox A6 - human (fragment)  
N:Alternate names: homeotic protein Hox 1B  
C:Species: *Homo sapiens* (man)  
C>Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 16-Aug-2004  
C:Accession: S15538  
R:Boncinelli, E.; Acampora, D.; Pannese, M.; d'Esposito, M.; Somma, R.; Gaudino, G.; Sto  
Genome 31, 745-756, 1989  
A:Title: Organization of human class I homeobox genes.  
A:Reference number: S15036; MUID:90215256; PMID:2576652  
A:Accession: S15538  
A:Status: not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-66 <BON>  
A:Cross-references: UNIPROT:P31267  
C:Genetics:  
A:Gene: GDB:HoxA6  
A:Cross-references: GDB:120648; OMIM:142951  
A:Map position: 7p15.3-7p15.3  
C:Superfamily: homeobox homology  
C:Keywords: DNA binding; homeobox; nucleus; transcription regulation  
F;2-58/Domain: homeobox homology <HOX>

Query Match 90.3%; Score 84; DB 2; Length 66;  
Best Local Similarity 93.8%; Pred. No. 6e-06;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RQIKWQNRRMKWFK 16  
| | | | | | | | | | | | | | | | | |  
Db 43 RQIKWQNRRMKWKK 58

RESULT 7  
S15536  
homeotic protein Hox A7 - human (fragment)  
N:Alternate names: homeotic protein Hox 1A  
C:Species: *Homo sapiens* (man)  
C>Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 16-Aug-2004  
C:Accession: S15536  
R:Boncinelli, E.; Acampora, D.; Pannese, M.; d'Esposito, M.; Somma, R.; Gaudino, G.; Sto  
Genome 31, 745-756, 1989  
A:Title: Organization of human class I homeobox genes.  
A:Reference number: S15036; MUID:90215256; PMID:2576652  
A:Accession: S15536  
A:Status: not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-66 <BON>  
A:Cross-references: UNIPROT:P31268  
C:Genetics:  
A:Gene: GDB:HoxA7  
A:Cross-references: GDB:120647; OMIM:142950  
A:Map position: 7p15.3-7p15.3  
C:Superfamily: homeobox homology  
C:Keywords: DNA binding; homeobox; nucleus; transcription regulation

F;2-58/Domain: homeobox homology <HOX>

Query Match 90.3%; Score 84; DB 2; Length 66;  
Best Local Similarity 93.8%; Pred. No. 6e-06;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RQIKWQNRRMKWFK 16  
|||  
Db 43 RQIKWQNRRMKWKK 58

#### RESULT 8

JC1161  
homeotic protein 3.4 - eastern newt (fragment)  
C;Species: Notophthalmus viridescens, Triturus viridescens (eastern newt)  
C;Date: 09-Oct-1992 #sequence\_revision 09-Oct-1992 #text\_change 16-Aug-2004  
C;Accession: JC1161  
R;Belleville, S.; Beauchemin, M.; Tremblay, M.; Noiseux, N.; Savard, P.  
Gene 114, 179-186, 1992  
A;Title: Homeobox-containing genes in the newt are organized in clusters similar to other

A;Reference number: JC1161; MUID:92290273; PMID:1351019

A;Accession: JC1161

A;Molecule type: DNA

A;Residues: 1-71 <BEL>

A;Cross-references: UNIPROT:P31262; GB:M84001

C;Genetics:

A;Gene: NVHBox-3.4

C;Superfamily: homeobox homology

C;Keywords: DNA binding; homeobox; nucleus; transcription regulation

F;5-61/Domain: homeobox homology <HOX>

Query Match 90.3%; Score 84; DB 2; Length 71;  
Best Local Similarity 93.8%; Pred. No. 6.4e-06;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RQIKWQNRRMKWFK 16  
|||  
Db 46 RQIKWQNRRMKWKK 61

#### RESULT 9

A60084  
homeotic protein Hox 3.4 - mouse (fragment)  
C;Species: Mus musculus (house mouse)  
C;Date: 20-Feb-1993 #sequence\_revision 20-Feb-1993 #text\_change 09-Jul-2004  
C;Accession: A60084  
R;Gaunt, S.J.; Coletta, P.L.; Pravtcheva, D.; Sharpe, P.T.  
Development 109, 329-339, 1990  
A;Title: Mouse Hox-3.4: homeobox sequence and embryonic expression patterns compared with

A;Reference number: A60084; MUID:90382249; PMID:1976088

A;Accession: A60084

A;Status: not compared with conceptual translation

A;Molecule type: DNA

A;Residues: 1-71 <GAU>

A;Cross-references: UNIPROT:P32043

C;Genetics:

A;Map position: 15

C;Superfamily: unassigned homeobox proteins; homeobox homology

C;Keywords: DNA binding; homeobox; nucleus; transcription regulation

F;5-61/Domain: homeobox homology <HOX>

Query Match 90.3%; Score 84; DB 2; Length 71;  
Best Local Similarity 93.8%; Pred. No. 6.4e-06;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RQIKWQNRRMKWFK 16  
|||  
Db 46 RQIKWQNRRMKWKK 61

#### RESULT 10

D34510  
homeotic protein H90 - honeybee (fragment)

C;Species: Apis mellifera (honeybee)

C;Date: 22-Jun-1990 #sequence\_revision 09-Oct-1992 #text\_change 16-Aug-2004

C;Accession: D34510

R;Walldorf, U.; Fleig, R.; Gehring, W.J.

Proc. Natl. Acad. Sci. U.S.A. 86, 9971-9975, 1989

A;Title: Comparison of homeobox-containing genes of the honeybee and Drosophila.

A;Reference number: A34510; MUID:9009384; PMID:2574865

A;Accession: D34510

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-74 <WAL>

A;Cross-references: UNIPROT:P15860; GB:M29493; NID:g155675; PIDN:AAA27728.1; PID:g155676

A;Note: the authors mistranslated the codons for residues 68-74

C;Superfamily: homeobox homology

C;Keywords: DNA binding; homeobox; nucleus; transcription regulation

F;9-65/Domain: homeobox homology <HOX>

Query Match 90.3%; Score 84; DB 2; Length 74;

Best Local Similarity 93.8%; Pred. No. 6.7e-06;

Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RQIKWQNRRMKWFK 16

|||  
Db 50 RQIKWQNRRMKWKK 65

#### RESULT 11

S58852  
homeotic protein Scr homolog - Junonia coenia (fragment)

N;Alternate names: sex combs reduced homeodomain protein

C;Species: Junonia coenia

C;Date: 19-Mar-1997 #sequence\_revision 29-Aug-1997 #text\_change 16-Aug-2004

C;Accession: S58852

R;Warren, R.W.; Nagy, L.; Sealeg, J.; Gates, J.; Carroll, S.

Nature 372, 458-461, 1994

A;Title: Evolution of homeotic gene regulation and function in flies and butterflies.

A;Reference number: S58850; MUID:95075456; PMID:7840822

A;Accession: S58852

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: mRNA

A;Residues: 1-75 <WAR>

A;Cross-references: UNIPROT:Q25209; EMBL:L42136; NID:g806495; PIDN:AAA68462.1; PID:g8064

A;Note: the nucleotide sequence was submitted to the EMBL Data Library, May 1995

C;Superfamily: homeobox homology

C;Keywords: DNA binding; homeobox; nucleus; transcription regulation

F;1-40/Domain: homeobox homology (fragment) <HOX>

Query Match 90.3%; Score 84; DB 2; Length 75;

Best Local Similarity 93.8%; Pred. No. 6.8e-06;

Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RQIKWQNRRMKWFK 16

|||  
Db 25 RQIKWQNRRMKWKK 40

#### RESULT 12

I51341  
homeo box protein - Atlantic salmon (fragment)

C;Species: Salmo salar (Atlantic salmon)

C;Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 16-Aug-2004

C;Accession: I51341

R;Fjose, A.; Molven, A.; Eiken, H.G.

Gene 62, 141-152, 1988

A;Title: Molecular cloning and characterization of homeobox-containing genes from Atlan

A;Reference number: I51341; MUID:88226009; PMID:2897318

A;Accession: I51341

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-75 <FJO>

A;Cross-references: UNIPROT:P09636; GB:M18903; NID:g213797; PIDN:AAA49559.1; PID:g213798

C;Superfamily: homeobox homology

C;Keywords: DNA binding; homeobox; nucleus; transcription regulation

F;2-58/Domain: homeobox homology <HOX>

Query Match 90.3%; Score 84; DB 2; Length 75;  
Best Local Similarity 93.8%; Pred. No. 6.8e-06;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RQIKWIFQNRMRMKWK 16  
| | | | | | | | | | | | | | | |  
Db 43 RQIKWIFQNRMRMKWK 58

#### RESULT 13

C43559  
homeotic protein R3 - rat (fragment)  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 01-Dec-1992 #sequence\_revision 01-Dec-1992 #text\_change 16-Aug-2004  
C;Accession: C43559  
R;Falzon, M.; Chung, S.Y.  
Development 103, 601-610, 1988  
A;Title: The expression of rat homeobox-containing genes is developmentally regulated and  
A;Reference number: A43559; MUID:89231502; PMID:2907739  
A;Accession: C43559  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-76 <PAL>  
A;Cross-references: UNIPROT:P18865; GB:M37567; NID:g204634; PIDN:AAA41343.1; PID:g204635  
C;Superfamily: homeobox homology  
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation  
F;20-76/Domain: homeobox homology (fragment) <HOX>

Query Match 90.3%; Score 84; DB 2; Length 76;  
Best Local Similarity 93.8%; Pred. No. 6.9e-06;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RQIKWIFQNRMRMKWK 16  
| | | | | | | | | | | | | | | |  
Db 53 RQIKWIFQNRMRMKWK 68

#### RESULT 14

I51342  
homeo box protein - Atlantic salmon (fragment)  
C;Species: Salmo salar (Atlantic salmon)  
C;Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 16-Aug-2004  
C;Accession: I51342  
R;Fjose, A.; Molven, A.; Eiken, H.G.  
Gene 62, 141-152, 1988  
A;Title: Molecular cloning and characterization of homeobox-containing genes from Atlantic  
A;Reference number: I51341; MUID:88226009; PMID:2897318  
A;Accession: I51342  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-78 <FJO>  
A;Cross-references: UNIPROT:P09637; GB:M18904; NID:g213799; PIDN:AAA49560.1; PID:g213800  
C;Superfamily: homeobox homology  
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation  
F;2-58/Domain: homeobox homology <HOX>

Query Match 90.3%; Score 84; DB 2; Length 78;  
Best Local Similarity 93.8%; Pred. No. 7.1e-06;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RQIKWIFQNRMRMKWK 16  
| | | | | | | | | | | | | | | |  
Db 43 RQIKWIFQNRMRMKWK 58

#### RESULT 15

B29585  
homeotic protein Hox 2.2 precursor - mouse (fragment)  
C;Species: Mus musculus (house mouse)  
C;Date: 15-Dec-1988 #sequence\_revision 30-Sep-1991 #text\_change 17-Oct-1997  
C;Accession: B29585

R;Lonai, P.; Arman, E.; Czosnek, H.; Ruddle, F.H.; Blatt, C.  
DNA 6, 409-418, 1987

A;Title: New murine homeoboxes: structure, chromosomal assignment, and differential expression.  
A;Reference number: A29585; MUID:88054465; PMID:2890503  
A;Accession: B29585  
A;Molecule type: DNA  
A;Residues: 1-81 <LON>  
A;Cross-references: GB:M18167  
A;Note: the authors translated the codon CAG for residue 69 as Glu  
C;Superfamily: unassigned homeobox proteins; homeobox homology  
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation  
F;4-60/Domain: homeobox homology <HOX>

Query Match 90.3%; Score 84; DB 2; Length 81;  
Best Local Similarity 93.8%; Pred. No. 7.3e-06;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RQIKWIFQNRMRMKWK 16  
| | | | | | | | | | | | | | | |  
Db 45 RQIKWIFQNRMRMKWK 60

Search completed: March 31, 2005, 00:31:37  
Job time : 3.09231 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2005, 00:19:36 ; Search time 8.92308 Seconds  
(without alignments)  
918.211 Million cell updates/sec

Title: US-10-049-822A-13  
Perfect score: 93  
Sequence: 1 RQIKWIFQNRRMKWFK 16

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt\_03.\*  
1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	84	90.3	33	2 Q86FU0	Q86fu0 drosophila
2	84	90.3	39	2 Q57368	Q57368 brachydanio
3	84	90.3	42	2 Q80WH6	Q80wh6 rattus sp.
4	84	90.3	43	2 Q57359	Q57359 brachydanio
5	84	90.3	44	2 Q618J3	Q618j3 lampetra ja
6	84	90.3	46	2 Q9PVR9	Q9pvr9 oryzias lat
7	84	90.3	47	2 Q61T83	Q61t83 oncopeltus
8	84	90.3	48	1 HXB6 XENLA	P31256 xenopus lae
9	84	90.3	49	1 HXA5 SHEEP	Q28599 ovis aries
10	84	90.3	50	2 Q14543	Q14543 homo sapien
11	84	90.3	51	2 Q23743	Q23743 ctenodrilus
12	84	90.3	51	2 Q26407	Q26407 ctenodrilus
13	84	90.3	51	2 Q72413	Q72413 ctenodrilus
14	84	90.3	51	2 Q7JPR7	Q7jpr7 ctenodrilus
15	84	90.3	57	2 Q9PVR8	Q9pvr8 oryzias lat
16	84	90.3	58	2 Q25208	Q25208 junonia coe
17	84	90.3	58	2 Q9Y188	Q9y188 priapulid c
18	84	90.3	58	2 Q57362	Q57362 brachydanio
19	84	90.3	59	2 Q8WRM9	Q8wr9 lithobius a
20	84	90.3	59	2 Q9NB42	Q9nb42 anopheles g
21	84	90.3	59	2 Q9PVR5	Q9pvr5 oryzias lat
22	84	90.3	60	2 Q77139	Q77139 archegozete
23	84	90.3	60	2 Q77143	Q77143 archegozete
24	84	90.3	60	2 Q80WH4	Q80wh4 rattus sp.
25	84	90.3	60	2 Q80WH7	Q80wh7 rattus sp.
26	84	90.3	60	2 Q8QGL2	Q8qgl2 petromyzon
27	84	90.3	60	2 Q8QGL3	Q8qgl3 petromyzon
28	84	90.3	60	2 Q8QGL5	Q8qgl5 petromyzon
29	84	90.3	60	2 Q8QGL6	Q8qgl6 petromyzon
30	84	90.3	60	2 Q8QGL7	Q8qgl7 petromyzon
31	84	90.3	60	2 Q8QGL8	Q8qgl8 petromyzon

32 84 90.3 61 2 Q27910 polyandroca  
33 84 90.3 63 2 Q77138 archegozete  
34 84 90.3 63 2 Q8MXB2 holopneuste  
35 84 90.3 66 2 Q57356 brachydanio  
36 84 90.3 69 2 Q9BMF7 haliotis as  
37 84 90.3 69 2 Q9U9T4 nereis vire  
38 84 90.3 70 2 Q967W5 folsomia ca  
39 84 90.3 70 2 Q801B4 latimeria m  
40 84 90.3 71 1 HXA7 SHEEP  
41 84 90.3 71 1 HXC5 NOTVI  
42 84 90.3 71 2 Q6EBC8 bugula turr  
43 84 90.3 71 2 Q9PVS1 oryzias lat  
44 84 90.3 71 2 Q9PVS3 oryzias lat  
45 84 90.3 73 2 Q86D93 spadella ce

#### ALIGNMENTS

RESULT 1  
Q86FU0  
ID Q86FU0 PRELIMINARY; PRT; 33 AA.  
AC Q86FU0;  
DT 01-JUN-2003 (TREMBlrel. 24, Created)  
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)  
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
DE Antennapedia complex (Fragment).  
GN Name=ANT-C;  
OS Drosophila pseudoobscura (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7237;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93292933; PubMed=8099892;  
RA Randazzo F.M., Seeger M.A., Huss C.A., Sweeney M.A., Cecil J.K.,  
RA Kaufman T.C.;  
RT "Structural changes in the antennapedia complex of Drosophila  
pseudoobscura";  
RL Genetics 134:319-330(1993).  
CC 1- SUBCELLULAR LOCATION: Nuclear (By similarity).  
DR HMBL; S63455; AAP13946.1; -.  
DR HSSP; P02833; IHOM.  
DR GO; GO:0005634; C:nucleus; IEA.  
DR GO; GO:0003700; F:transcription factor activity; IEA.  
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
DR InterPro; IPR001356; Homeobox.  
DR InterPro; IPR009057; Homeodomain\_like.  
DR PRINTS; PR00024; HOMEBOX.  
DR PRODom; PD000010; Homeobox; 1.  
DR PROSITE; PS00071; HOMEBOX 2; 1.  
KW DNA-binding; Homeobox; Nuclear protein.  
FT NON\_TER 1  
FT NON\_TER 33  
SQ SEQUENCE 33 AA; 3963 MW; D78E37ED81FD45DF CRC64;

Query Match 90.3%; Score 84; DB 2; Length 33;  
Best Local Similarity 93.8%; Pred. No. 8.4e-06;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 RQIKWIFQNRRMKWFK 16  
| | | | | | | | | | | | | | | | | |  
Db 7 RQIKWIFQNRRMKWFK 22

RESULT 2  
Q57368  
ID Q57368 PRELIMINARY; PRT; 39 AA.  
AC Q57368;  
DT 01-JUN-1998 (TREMBlrel. 06, Created)  
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)  
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)

DE Hoxc5 protein (Fragment).  
GN Name=hoxc5a; Synonyms=hoxc5;  
OS Brachydanio rerio (Zebrafish) (Danio rerio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
OX NCBI\_TaxID=7955;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Prince V.E., Joly L., Ekker M., Ho R.K.;  
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.  
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).  
DR EMBL; Y14539; CAA74874.1; -.  
DR HSSP; P02833; IHOM.  
DR ZFIN; ZDB-GENE-980526-533; hoxc5a.  
DR GO; GO:0005634; C:nucleus; IEA.  
DR GO; GO:0003700; F:transcription factor activity; IEA.  
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
DR InterPro; IPR001356; Homeobox.  
DR InterPro; IPR009057; Homeodomain-like.  
DR InterPro; IPR000047; HTH\_lambrepresr.  
DR Pfam; PF00046; Homeobox; 1.  
DR PRINTS; PR00024; HOMEBOX.  
DR PRODOM; PD000010; Homeobox; 1.  
DR SMART; SM00389; HOX; 1.  
DR PROSITE; PS00027; HOMEBOX\_1; 1.  
DR PROSITE; PS50071; HOMEBOX\_2; 1.  
KW DNA-binding; Homeobox; Nuclear protein.  
FT NON TER 1  
SQ SEQUENCE 39 AA; 4827 MW; 592A0FEC12B58860 CRC64;  
  
Query Match 90.3%; Score 84; DB 2; Length 39;  
Best Local Similarity 93.8%; Pred. No. 1.e-05;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 1 RQIKWIFQNRMRMKWPK 16  
Db 14 RQIKWIFQNRMRMKWK 29  
  
RESULT 3  
Q80WH6 PRELIMINARY; PRT; 42 AA.  
AC Q80WH6;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Hox-A|Hox-1 (Fragment).  
GN Name=Hox-A|Hox-1;  
OS Rattus sp.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10118;  
RN [1]  
RP SEQUENCE FROM N.A. PubMed=7702549;  
RX MEDLINE=95217128; PubMed=7702549;  
RA Sakoyama Y., Mizuta I., Ogasawara N., Yoshikawa H.;  
RT "Cloning of rat homeobox genes."  
RL Biochem. Genet. 32:351-360(1994).  
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).  
DR EMBL; S76290; AAP31864.1; -.  
DR HSSP; P02833; IHOM.  
DR GO; GO:0005634; C:nucleus; IEA.  
DR GO; GO:0003700; F:transcription factor activity; IEA.  
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
DR InterPro; IPR001356; Homeobox.  
DR InterPro; IPR009057; Homeodomain-like.  
DR InterPro; IPR000047; HTH\_lambrepresr.  
DR Pfam; PF00046; Homeobox; 1.  
DR PRINTS; PR00024; HOMEBOX.  
DR PRODOM; PD000010; Homeobox; 1.

DR SMART; SM00389; HOX; 1.  
DR PROSITE; PS00027; HOMEBOX\_1; 1.  
DR PROSITE; PS50071; HOMEBOX\_2; 1.  
KW DNA-binding; Homeobox; Nuclear protein.  
FT NON TER 1  
FT NON TER 42  
SQ SEQUENCE 42 AA; 5494 MW; 38B5153B92216FE9 CRC64;  
  
Query Match 90.3%; Score 84; DB 2; Length 42;  
Best Local Similarity 93.8%; Pred. No. 1.e-05;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 1 RQIKWIFQNRMRMKWPK 16  
Db 25 RQIKWIFQNRMRMKWK 40  
  
RESULT 4  
O57359 PRELIMINARY; PRT; 43 AA.  
AC O57359;  
DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Hoxa5 protein (Fragment).  
GN Name=hoxb5b; Synonyms=hoxa5;  
OS Brachydanio rerio (Zebrafish) (Danio rerio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
OX NCBI\_TaxID=7955;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Prince V.E., Joly L., Ekker M., Ho R.K.;  
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.  
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).  
DR EMBL; Y14526; CAA74861.1; -.  
DR HSSP; P02833; IHOM.  
DR ZFIN; ZDB-GENE-000823-6; hoxb5b.  
DR GO; GO:0005634; C:nucleus; IEA.  
DR GO; GO:0003700; F:transcription factor activity; IEA.  
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
DR InterPro; IPR001356; Homeobox.  
DR InterPro; IPR009057; Homeodomain-like.  
DR Pfam; PF00046; Homeobox; 1.  
DR PRINTS; PR00024; HOMEBOX.  
DR PRODOM; PD000010; Homeobox; 1.  
DR PROSITE; PS00027; HOMEBOX\_1; 1.  
DR PROSITE; PS50071; HOMEBOX\_2; 1.  
KW DNA-binding; Homeobox; Nuclear protein.  
FT NON TER 1  
SQ SEQUENCE 43 AA; 5050 MW; 53034C37F3DFA596 CRC64;  
  
Query Match 90.3%; Score 84; DB 2; Length 43;  
Best Local Similarity 93.8%; Pred. No. 1.e-05;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 1 RQIKWIFQNRMRMKWPK 16  
Db 10 RQIKWIFQNRMRMKWK 25  
  
RESULT 5  
Q6L8J3 PRELIMINARY; PRT; 44 AA.  
AC Q6L8J3;  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE LjHox7m Homeobox (Fragment).  
GN Name=LjHox7m;  
OS Lampetra japonica (Japanese lamprey) (Entosphenus japonicus).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;



OC Petromyzontiformes; Petromyzontidae; Lethenteron.  
RN NCBI\_TaxID=94989;  
RP SEQUENCE FROM N.A.  
RA Takio Y., Kuraku S., Kuratani S.;  
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.  
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).  
DR EMBL; AB125272; BAD2526.1; -.  
DR HSSP; P02833; 1AHD.  
DR GO; GO:0005634; C:nucleus; IEA.  
DR GO; GO:0003700; F:transcription factor activity; IEA.  
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
DR InterPro; IPR001356; Homeobox.  
DR InterPro; IPR009057; Homeobox-like.  
DR InterPro; IPR000047; HTH lambrapress.  
DR Pfam; PF00046; Homeobox\_1.  
DR PRINTS; PR00024; HOMEBOX.  
DR PRODOM; PD000010; Homeobox; 1.  
DR SMART; SM00389; HOX; 1.  
DR PROSITE; PS00027; HOMEBOX 1; 1.  
DR PROSITE; PS00071; HOMEBOX\_2; 1.  
KW DNA-binding; Homeobox; Nuclear protein.  
FT NON\_TER 1 1  
FT NON\_TER 44 44  
SQ SEQUENCE 44 AA; 5668 MW; B608B0F91C51ED11 CRC64;  
  
Query Match 90.3%; Score 84; DB 2; Length 44;  
Best Local Similarity 93.8%; Pred. No. 1.1e-05;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 1 RQIKWIFQNRMRKWK 16  
Db 23 RQIKWIFQNRMRKWK 38  
  
RESULT 6  
Q9PVR9 PRELIMINARY; PRT; 46 AA.  
AC Q9PVR9;  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
DE HOXC5A (Fragment).  
GN Name=hoxc5a;  
OS Oryzias latipes (Medaka fish) (Japanese ricefish).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;  
OC Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.  
OX NCBI\_TaxID=8090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Kondo S., Naruse K., Shima A.;  
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).  
DR EMBL; AB026960; BAA86243.1; -.  
DR HSSP; P02833; 9ANT.  
DR GO; GO:0005634; C:nucleus; IEA.  
DR GO; GO:0003700; F:transcription factor activity; IEA.  
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
DR InterPro; IPR001356; Homeobox.  
DR InterPro; IPR009057; Homeobox-like.  
DR Pfam; PF00046; Homeobox; 1.  
DR PRINTS; PR00024; HOMEBOX.  
DR PRODOM; PD000010; Homeobox; 1.  
DR SMART; SM00389; HOX; 1.  
DR PROSITE; PS00027; HOMEBOX 1; 1.  
DR PROSITE; PS00071; HOMEBOX\_2; 1.  
KW DNA-binding; Homeobox; Nuclear protein.  
FT NON\_TER 1 1  
FT NON\_TER 46 46  
SQ SEQUENCE 46 AA; 5955 MW; 6039999ED4294DD3 CRC64;

Query Match 90.3%; Score 84; DB 2; Length 46;  
Best Local Similarity 93.8%; Pred. No. 1.2e-05;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 1 RQIKWIFQNRMRKWK 16  
Db 23 RQIKWIFQNRMRKWK 38  
  
RESULT 7  
Q6IT83 PRELIMINARY; PRT; 47 AA.  
AC Q6IT83;  
DT 05-JUL-2004 (TREMBlrel. 27, Created)  
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
DE Antennapedia (Fragment).  
OS Oncopeltus fasciatus (Milkweed bug).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Paraneoptera; Hemiptera; Euhemiptera; Heteroptera;  
OC Panheteroptera; Pentatomomorpha; Lygaeoidea; Lygaeinae;  
OC Oncopeltus.  
OX NCBI\_TaxID=7536;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX PubMed=15342481;  
RA Liu P.Z., Kaufman T.C.;  
RT "Kruppel is a gap gene in the intermediate germband insect *Oncopeltus*  
RT fasciatus and is required for development of both blastoderm and  
RT germband-derived segments.";  
RL Development 131:4567-4579(2004).  
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).  
DR EMBL; AY627363; AAT44525.1; -.  
DR HSSP; P02833; 1AHD.  
DR GO; GO:0005634; C:nucleus; IEA.  
DR GO; GO:0003700; F:transcription factor activity; IEA.  
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
DR InterPro; IPR001356; Homeobox.  
DR InterPro; IPR009057; Homeobox-like.  
DR Pfam; PF00046; Homeobox; 1.  
DR PRINTS; PR00024; HOMEBOX.  
DR PRODOM; PD000010; Homeobox; 1.  
DR PROSITE; PS00071; HOMEBOX\_2; 1.  
KW DNA-binding; Homeobox; Nuclear protein.  
FT NON\_TER 1 1  
SQ SEQUENCE 47 AA; 5441 MW; ABBAE19564222859 CRC64;  
  
Query Match 90.3%; Score 84; DB 2; Length 47;  
Best Local Similarity 93.8%; Pred. No. 1.2e-05;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 1 RQIKWIFQNRMRKWK 16  
Db 9 RQIKWIFQNRMRKWK 24  
  
RESULT 8  
HXB6 XENLA STANDARD; PRT; 48 AA.  
AC P31256;  
DT 01-JUL-1993 (Rel. 26, Created)  
DT 01-JUL-1993 (Rel. 26, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Homeobox protein Hox-B6 (XlHox-2.2) (Fragment).  
GN Name=HXB6; Synonym=XLHox-2.2;  
OS *Xenopus laevis* (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;  
OC Xenopodinae; Xenopus.  
OX NCBI\_TaxID=8355;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93043517; PubMed=1384809;

RA Leroy P., de Robertis E.M.;  
RT "Effects of lithium chloride and retinoic acid on the expression of  
RL genes from the *Xenopus laevis* Hox 2 complex.";   
RL Dyn. Dev. 194;21-32(1992).  
CC -!- FUNCTION: Sequence-specific transcription factor which is part of  
CC a developmental regulatory system that provides cells with  
CC specific positional identities on the anterior-posterior axis.  
CC -!- SUBCELLULAR LOCATION: Nuclear.  
CC -!- SIMILARITY: Belongs to the Antp homeobox family.  
CC -!- SIMILARITY: Contains 1 homeobox domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; M91587; AAA49750.1; -.  
DR PIR; I51439; I51439.  
DR HSSP; P02833; IHOM.  
DR InterPro; IPR001827; Antennapedia.  
DR InterPro; IPR001356; Homeobox.  
DR InterPro; IPR009057; Homeobox\_like.  
DR Pfam; PF00046; Homeobox; 1.  
DR PRINTS; PR00024; HOMEBOX.  
DR PRODOM; PD000010; Homeobox; 1.  
DR SMART; SM00389; HOX; 1.  
DR PROSITE; PS00027; HOMEBOX 1; 1.  
DR PROSITE; PS50071; HOMEBOX 2; 1.  
DR Developmental protein; DNA-binding; Homeobox; Nuclear protein;  
KW Transcription regulation.  
FT NON TER 1 1  
FT DNA BIND <1 29 Homeobox.  
FT NON TER 1 1  
SQ SEQUENCE 48 AA; 5716 MW; BC39E36822EDDD2A CRC64;  
  
Query Match 90.3%; Score 84; DB 1; Length 48;  
Best Local Similarity 93.8%; Pred. No. 1.2e-05;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 1 RQIKWFFQNRMRKWK 16  
Db 12 RQIKWFFQNRMRKWK 27  
  
RESULT 9  
HXAS SHEEP  
ID \_HXAS5 SHEEP STANDARD; PRT; 49 AA.  
AC Q28599;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Homeobox protein Hox-A5 (Fragment).  
CN Name=HOXA5; Synonyms=HOXA-5;  
OS Ovis aries (Sheep).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Caprinae; Ovis.  
OX NCBI\_TaxID=9940;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Roche P.J.;  
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: Sequence-specific transcription factor which is part of  
CC a developmental regulatory system that provides cells with  
CC specific positional identities on the anterior-posterior axis.  
CC Also binds to its own promoter. Binds specifically to the motif  
CC 5'-CYNATTA[GG]Y-3'.  
CC -!- SUBCELLULAR LOCATION: Nuclear.  
CC -!- SIMILARITY: Belongs to the Antp homeobox family.  
CC -!- SIMILARITY: Contains 1 homeobox domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; U61978; AAB04754.1; -.  
DR HSSP; P02833; IHOM.  
DR InterPro; IPR001356; Homeobox.  
DR InterPro; IPR009057; Homeobox\_like.  
DR Pfam; PF00046; Homeobox; 1.  
DR PRINTS; PR00024; HOMEBOX.  
DR PRODOM; PD000010; Homeobox; 1.  
DR SMART; SM00389; HOX; 1.  
DR PROSITE; PS00027; HOMEBOX 1; 1.  
DR PROSITE; PS50071; HOMEBOX 2; 1.  
DR Developmental protein; DNA-binding; Homeobox; Nuclear protein;  
KW Transcription regulation.  
FT NON TER 1 1  
FT DNA BIND <1 49 Homeobox.  
FT NON TER 49 49  
SQ SEQUENCE 49 AA; 6331 MW; 1EE702315E7C099B CRC64;  
  
Query Match 90.3%; Score 84; DB 1; Length 49;  
Best Local Similarity 93.8%; Pred. No. 1.3e-05;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 1 RQIKWFFQNRMRKWK 16  
Db 32 RQIKWFFQNRMRKWK 47  
  
RESULT 10  
Q14543  
ID Q14543 PRELIMINARY; PRT; 50 AA.  
AC Q14543;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Homeobox protein (50 AA) (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RA Gilman J.G.;  
RL Submitted (APR-1990) to the EMBL/GenBank/DBJ databases.  
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).  
CC EMBL; X52402; CAA36647.1; -.  
DR PIR; S10907; S10907.  
DR HSSP; P02833; 9ANT.  
DR GO; GO:0005634; C:nucleus; IEA.  
DR GO; GO:0003700; F:transcription factor activity; IEA.  
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
DR InterPro; IPR001356; Homeobox.  
DR InterPro; IPR009057; Homeobox\_like.  
DR Pfam; PF00046; Homeobox; 1.  
DR PRINTS; PR00024; HOMEBOX.  
DR PRODOM; PD000010; Homeobox; 1.  
DR SMART; SM00389; HOX; 1.  
DR PROSITE; PS00027; HOMEBOX 1; 1.  
DR PROSITE; PS50071; HOMEBOX 2; 1.  
KW DNA-binding; Homeobox; Nuclear protein.  
FT NON TER 1 1  
SQ SEQUENCE 50 AA; 6411 MW; 9E90625E455ED582 CRC64;  
  
Query Match 90.3%; Score 84; DB 2; Length 50;  
Best Local Similarity 93.8%; Pred. No. 1.3e-05;

```
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ROIKIWFQNRMRKWK 16
Db 25 ROIKIWFQNRMRKWK 40

RESULT 11
Q23743 ID Q23743 PRELIMINARY; PRT; 51 AA.
AC Q23743;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Deformed ortholog homeobox (Fragment).
GN Name=CTS-Dfd;
OS Ctenodrilus serratus.
OC Eukaryota; Metazoa; Annelida; Polychaeta; Palpata; Canaliculipalata;
OC Terebellida; Ctenodrilidae; Ctenodrilus.
OX NCBI_TaxID=40316;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94356262; PubMed=7915607; DOI=10.1006/mpev.1994.1017;
RA "A PCR-based survey of homeobox genes in Ctenodrilus serratus
RT (Annelida: Polychaeta).";
RL Mol. Phylogenet. Evol. 3:146-158(1994).
CC -|- SUBCELLULAR LOCATION: Nuclear (By similarity).
DR EMBL; S76416; AAC31775.1; -.
DR HSSP; P02833; 9ANT.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR InterPro; IPR009057; HTH lambdarepressor.
DR Pfam; PF00046; Homeobox_1.
DR PRINTS; PR00024; HOMEBOX.
DR PRINTS; PR00031; HTHREPRESSR.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS00071; HOMEBOX_2; 1.
KW DNA-binding; Homeobox; Nuclear protein.
FT NON_TER 1 51
FT NON_TER 1 51
SQ SEQUENCE 51 AA; 6533 MW; 9EDB50C927FBCBD5 CRC64;

Query Match 90.3%; Score 84; DB 2; Length 51;
Best Local Similarity 93.8%; Pred. No. 1.3e-05;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ROIKIWFQNRMRKWK 16
Db 23 ROIKIWFQNRMRKWK 38

RESULT 13
Q27413 ID Q27413 PRELIMINARY; PRT; 51 AA.
AC Q27413;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Lox5 ortholog homeobox (Fragment).
GN Name=CTS-Lox5;
OS Ctenodrilus serratus.
OC Eukaryota; Metazoa; Annelida; Polychaeta; Palpata; Canaliculipalata;
OC Terebellida; Ctenodrilidae; Ctenodrilus.
OX NCBI_TaxID=40316;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94356262; PubMed=7915607; DOI=10.1006/mpev.1994.1017;
RA "A PCR-based survey of homeobox genes in Ctenodrilus serratus
RT (Annelida: Polychaeta).";
RL Mol. Phylogenet. Evol. 3:146-158(1994).
CC -|- SUBCELLULAR LOCATION: Nuclear (By similarity).
DR EMBL; U26629; AAC46851.1; -.
DR HSSP; P02833; 9ANT.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR InterPro; IPR009057; HTH lambdarepressor.
DR Pfam; PF00046; Homeobox_1.
DR PRINTS; PR00024; HOMEBOX.
DR PRINTS; PR00031; HTHREPRESSR.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS00071; HOMEBOX_2; 1.
KW DNA-binding; Homeobox; Nuclear protein.
FT NON_TER 1 51
FT NON_TER 1 51
SQ SEQUENCE 51 AA; 6533 MW; 9EDB50C927FBCBD5 CRC64;

Query Match 90.3%; Score 84; DB 2; Length 51;
Best Local Similarity 93.8%; Pred. No. 1.3e-05;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ROIKIWFQNRMRKWK 16
Db 23 ROIKIWFQNRMRKWK 38

RESULT 12
Q26407 ID Q26407 PRELIMINARY; PRT; 51 AA.
AC Q26407;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE CTS-Dfd protein (Fragment).
GN Name=CTS-Dfd;
OS Ctenodrilus serratus.
OC Eukaryota; Metazoa; Annelida; Polychaeta; Palpata; Canaliculipalata;
OC Terebellida; Ctenodrilidae; Ctenodrilus.
OX NCBI_TaxID=40316;
```

```

RN RP SEQUENCE FROM N.A.
RX MEDLINE=94356262; PubMed=7915607; DOI=10.1006/mpev.1994.1017;
RA "A PCR-based survey of homeobox genes in Ctenodrilus serratus
RT (Annelida: Polychaeta).";
RL Mol. Phylogenet. Evol. 3:146-158(1994).
CC -|- SUBCELLULAR LOCATION: Nuclear (By similarity).
DR EMBL; S76416; AAC31775.1; -.
DR HSSP; P02833; 9ANT.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR InterPro; IPR009057; HTH lambdarepressor.
DR Pfam; PF00046; Homeobox_1.
DR PRINTS; PR00024; HOMEBOX.
DR PRINTS; PR00031; HTHREPRESSR.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS00071; HOMEBOX_2; 1.
KW DNA-binding; Homeobox; Nuclear protein.
FT NON_TER 51 51
FT NON_TER 51 51
SQ SEQUENCE 51 AA; 6533 MW; 9EDB50C927FBCBD5 CRC64;

Query Match 90.3%; Score 84; DB 2; Length 51;
Best Local Similarity 93.8%; Pred. No. 1.3e-05;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ROIKIWFQNRMRKWK 16
Db 23 ROIKIWFQNRMRKWK 38

RESULT 13
Q27413 ID Q27413 PRELIMINARY; PRT; 51 AA.
AC Q27413;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Lox5 ortholog homeobox (Fragment).
GN Name=CTS-Lox5;
OS Ctenodrilus serratus.
OC Eukaryota; Metazoa; Annelida; Polychaeta; Palpata; Canaliculipalata;
OC Terebellida; Ctenodrilidae; Ctenodrilus.
OX NCBI_TaxID=40316;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94356262; PubMed=7915607; DOI=10.1006/mpev.1994.1017;
RA "A PCR-based survey of homeobox genes in Ctenodrilus serratus
RT (Annelida: Polychaeta).";
RL Mol. Phylogenet. Evol. 3:146-158(1994).
CC -|- SUBCELLULAR LOCATION: Nuclear (By similarity).
DR EMBL; U26629; AAC46851.1; -.
DR HSSP; P02833; 9ANT.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR InterPro; IPR009057; HTH lambdarepressor.
DR Pfam; PF00046; Homeobox_1.
DR PRINTS; PR00024; HOMEBOX.
DR PRINTS; PR00031; HTHREPRESSR.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS00071; HOMEBOX_2; 1.
KW DNA-binding; Homeobox; Nuclear protein.
FT NON_TER 51 51
FT NON_TER 51 51
SQ SEQUENCE 51 AA; 6533 MW; 9EDB50C927FBCBD5 CRC64;

Query Match 90.3%; Score 84; DB 2; Length 51;
Best Local Similarity 93.8%; Pred. No. 1.3e-05;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ROIKIWFQNRMRKWK 16
Db 23 ROIKIWFQNRMRKWK 38

RESULT 12
Q26407 ID Q26407 PRELIMINARY; PRT; 51 AA.
AC Q26407;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE CTS-Dfd protein (Fragment).
GN Name=CTS-Dfd;
OS Ctenodrilus serratus.
OC Eukaryota; Metazoa; Annelida; Polychaeta; Palpata; Canaliculipalata;
OC Terebellida; Ctenodrilidae; Ctenodrilus.
OX NCBI_TaxID=40316;
```

```
DR PROSITE; PS00071; HOMEBOX 2; 1.
KW DNA-binding; Homeobox; Nuclear protein.
FT NON_TER 1
FT NON_TER 51
SQ SEQUENCE 51 AA; 6278 MW; 88C8F65161E94A22 CRC64;

Query Match          90.3%; Score 84; DB 2; Length 51;
Best Local Similarity 93.8%; Pred. No. 1.3e-05;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RQIKIWQNRRMKWFK 16
Db 23 RQIKIWQNRRMKWKK 38

RESULT 14
QJUPR7 PRELIMINARY; PRT; 51 AA.
AC Q7JUPR7;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE CTS-Lox5 protein (Fragment).
GN Name=CTS-Lox5;
OS Ctenodrilus serratus.
OC Eukaryota; Metazoa; Annelida; Polychaeta; Palpata; Canalipalpata;
OC Terebellida; Ctenodrilidae; Ctenodrilus.
OX NCBI_TaxID=40316;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94355262; PubMed=7915607; DOI=10.1006/mpev.1994.1017;
RA Dick M.H., Buss L.W.;
RT "A PCR-based survey of homeobox genes in Ctenodrilus serratus
  (Annelida: Polychaeta).";
RL Mol. Phylogenet. Evol. 3:146-158(1994).
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
DR EMBL; S76226; AAB31777.1; -.
DR HSSP; P83949; 1B81.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR InterPro; IPR009057; Homeodomain like.
DR InterPro; IPR000047; HTH_lambrepresr.
DR Pfam; PF00046; Homeobox_1.
DR PRINTS; PR00024; HOMEBOX.
DR PRINTS; PR00031; HTHREPRESSR.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX 1; 1.
DR PROSITE; PS50071; HOMEBOX_2; 1.
KW DNA-binding; Homeobox; Nuclear protein.
FT NON_TER 51
FT NON_TER 51
SQ SEQUENCE 51 AA; 6278 MW; 88C8F65161E94A22 CRC64;

Query Match          90.3%; Score 84; DB 2; Length 51;
Best Local Similarity 93.8%; Pred. No. 1.3e-05;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RQIKIWQNRRMKWFK 16
Db 23 RQIKIWQNRRMKWKK 38

RESULT 15
Q9PVR8 PRELIMINARY; PRT; 57 AA.
AC Q9PVR8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE HOXA5A (Fragment).
GN Name=hoxa5a;
```

```
OS Oryzias latipes (Medaka fish) (Japanese ricefish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
OX NCBI_TaxID=8090;
[1]
RP SEQUENCE FROM N.A.
RA Kondo S., Naruse K., Shima A.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
DR EMBL; AB026961; BAA86244.1; -.
DR HSSP; P02833; 1HOM.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR InterPro; IPR009057; Homeodomain like.
DR Pfam; PF00046; Homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX 1; 1.
DR PROSITE; PS50071; HOMEBOX 2; 1.
KW DNA-binding; Homeobox; Nuclear protein.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 57 AA; 6891 MW; 54A6430320F68C04 CRC64;

Query Match          90.3%; Score 84; DB 2; Length 57;
Best Local Similarity 93.8%; Pred. No. 1.5e-05;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RQIKIWQNRRMKWFK 16
Db 23 RQIKIWQNRRMKWKK 38

Search completed: March 31, 2005, 00:30:55
Job time : 11.9231 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2005, 00:25:06 ; Search time 5.28846 Seconds  
(without alignments)  
688.690 Million cell updates/sec

Title: US-10-049-822A-12

Perfect score: 58

Sequence: 1 YGRKKRQRRR 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1407402 seqs, 331100923 residues

Total number of hits satisfying chosen parameters: 1407402

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:\*  
1: /cgn2\_6/ptodata/2/pubaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/2/pubaa/US10\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/2/pubaa/US11\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/2/pubaa/US60\_NEW\_PUB.pep.\*  
20: /cgn2\_6/ptodata/2/pubaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	58	100.0	11	9 US-09-779-791A-1	Sequence 1, Appli
2	58	100.0	11	9 US-09-780-070-37	Sequence 37, Appl
3	58	100.0	11	9 US-09-815-108-9	Sequence 9, Appli
4	58	100.0	11	9 US-09-886-404-13	Sequence 13, Appl
5	58	100.0	11	9 US-09-805-805-8	Sequence 24, Appli
6	58	100.0	11	9 US-09-821-821-24	Sequence 8, Appli
7	58	100.0	11	9 US-09-895-943-13	Sequence 13, Appl
8	58	100.0	11	9 US-09-825-414-91	Sequence 91, Appl
9	58	100.0	11	9 US-09-867-274-23	Sequence 23, Appl
10	58	100.0	11	9 US-09-953-866-24	Sequence 24, Appl
11	58	100.0	11	9 US-09-928-175-25	Sequence 25, Appl
12	58	100.0	11	9 US-09-995-542-13	Sequence 13, Appl
13	58	100.0	11	9 US-09-927-850-18	Sequence 18, Appl

14	58	100.0	11	9 US-09-984-183-4	Sequence 4, Appli
15	58	100.0	11	9 US-09-949-196-15	Sequence 15, Appl
16	58	100.0	11	9 US-09-948-018-22	Sequence 22, Appl
17	58	100.0	11	9 US-09-995-515-14	Sequence 14, Appl
18	58	100.0	11	9 US-09-895-593-13	Sequence 13, Appl
19	58	100.0	11	9 US-09-896-738-20	Sequence 20, Appl
20	58	100.0	11	9 US-09-981-286A-2	Sequence 2, Appli
21	58	100.0	11	10 US-09-134-793-2	Sequence 2, Appli
22	58	100.0	11	10 US-09-882-291-43	Sequence 43, Appl
23	58	100.0	11	10 US-09-775-052-2	Sequence 2, Appli
24	58	100.0	11	10 US-09-847-946A-124	Sequence 124, App
25	58	100.0	11	10 US-09-876-904A-5	Sequence 5, Appli
26	58	100.0	11	10 US-09-876-904A-10	Sequence 10, Appl
27	58	100.0	11	10 US-09-845-612B-20	Sequence 20, Appl
28	58	100.0	11	10 US-09-997-465B-2	Sequence 2, Appli
29	58	100.0	11	10 US-09-911-261A-18	Sequence 18, Appl
30	58	100.0	11	10 US-09-798-053-10	Sequence 10, Appl
31	58	100.0	11	12 US-09-779-791A-1	Sequence 1, Appli
32	58	100.0	11	13 US-10-007-363-5	Sequence 5, Appli
33	58	100.0	11	14 US-10-139-833-18	Sequence 18, Appl
34	58	100.0	11	14 US-10-142-356-23	Sequence 23, Appl
35	58	100.0	11	14 US-10-156-424A-14	Sequence 14, Appl
36	58	100.0	11	14 US-10-208-374-2	Sequence 2, Appli
37	58	100.0	11	14 US-10-156-932-42	Sequence 42, Appl
38	58	100.0	11	14 US-10-077-555-4	Sequence 4, Appli
39	58	100.0	11	14 US-10-057-408-18	Sequence 18, Appl
40	58	100.0	11	14 US-10-229-915-28	Sequence 28, Appl
41	58	100.0	11	14 US-10-229-584-9	Sequence 9, Appli
42	58	100.0	11	14 US-10-156-487A-7	Sequence 7, Appli
43	58	100.0	11	14 US-10-193-616-15	Sequence 15, Appl
44	58	100.0	11	14 US-10-251-947-9	Sequence 9, Appli
45	58	100.0	11	14	

#### ALIGNMENTS

#### RESULT 1

US-09-779-791A-1  
; Sequence 1, Application US/09779791A  
; Publication No. US20010044417A1  
; GENERAL INFORMATION:  
; APPLICANT: Mirus Corporation  
; APPLICANT: Wolff, Jon A  
; APPLICANT: Monahan, Sean D  
; APPLICANT: Budker, Vladimir G  
; APPLICANT: Slattum, Paul M  
; APPLICANT: Rozema, David B  
; TITLE OF INVENTION: A Compound Containing a Labile Disulfide Bond  
; FILE REFERENCE: Mirus.006.03  
; CURRENT APPLICATION NUMBER: US/09/779,791A  
; CURRENT FILING DATE: 2001-02-08  
; PRIOR APPLICATION NUMBER: 09/312,351  
; PRIOR FILING DATE: 1999-05-14  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: Human immunodeficiency virus  
US-09-779-791A-1

Query Match 100.0%; Score 58; DB 9; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.018;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YGRKKRQRRR 11

Db 1 YGRKKRQRRR 11

#### RESULT 2

US-09-780-070-37

```

: Sequence 37, Application US/09780070
: Patent No. US2020009752A1
: GENERAL INFORMATION:
: APPLICANT: Burke, James
: APPLICANT: Strittmatter, Warren
: APPLICANT: Nagai, Yoshitaka
: TITLE OF INVENTION: COMPOUNDS THAT
: TITLE OF INVENTION: AND METHODS OF
: FILE REFERENCE: 5405.242
: CURRENT APPLICATION NUMBER: US/09/71
: CURRENT FILING DATE: 2001-02-09
: PRIOR APPLICATION NUMBER: 60/189,78
: PRIOR FILING DATE: 2000-03-16
: NUMBER OF SEQ ID NOS: 40
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 37
: LENGTH: 11
: TYPE: PRT
: ORGANISM: Human immunodeficiency v
: US-09-780-070-37

```

Query Match 100.0%; Score 58; DB 9; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.018;  
Matches 11; Conservative 0; Mismatches 0; Indels

Qy 1 YGRKRRQRRR 11  
Db 1 YGRKRRQRRR 11

### RESULT 3

```

US-09-815-108-9
; Sequence 9, Application US/09815108
; Patent No. US20020009776A1
; GENERAL INFORMATION:
; APPLICANT: Saris, Christiaan M.
; APPLICANT: Sharon, Wu X.
; APPLICANT: Xia, Min
; APPLICANT: Boone, Thomas Charles
; APPLICANT: Covey, Todd
; TITLE OF INVENTION: Fibroblast Growth Factor
; FILE REFERENCE: 99-513-A
; CURRENT APPLICATION NUMBER: US/09/09815108
; CURRENT FILING DATE: 2001-03-22
; PRIORITY APPLICATION NUMBER: 60/191,333
; PRIOR FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus
US-09-815-108-9

```

Query Match 100.0%; Score 58; DB 9; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.018;  
Matches 11; Conservative 0; Mismatches 0; Indels

Qy 1 YGRKRRQRRR 11  
|||  
Db 1 YGRKRRRORRR 11

## RESULT 4

```

US-09-886-404-13
; Sequence 13, Application US/09886404
; Patent No. US20020037524A1
; GENERAL INFORMATION:
; APPLICANT: Medlock, Eugene
; APPLICANT: Yen, Richard
; APPLICANT: Ellibiger, Scott M.
; APPLICANT: Elliott, Gary S.

```

```

; APPLICANT: Nguyen, Hung Q.
; APPLICANT: Jing, Shudian
; TITLE OF INVENTION: IL-17 Like Molecules and Uses Thereof
; FILE REFERENCE: 01017/37128B
; CURRENT APPLICATION NUMBER: US/09/886,404
; CURRENT FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: 09/810,384
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/266,159
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/213,125
; PRIOR FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 13
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Peptide of HIV
; OTHER INFORMATION: TAT protein
US-09-886-404-13

```

Query Match	100.0%;	Score 58;	DB 9;	Length 11;
Best Local Similarity	100.0%;	Pred. No. 0.018;		
Matches 11; Conservative	0;	Mismatches	0;	Gaps 0;

Qy 1 YGRKKRRQRRR 11  
|||  
db 1 YGRKKRRQRRR 11

## RESULT 5

```

US-09-805-805-8
; Sequence 8, Application US/09805805
; Patent No. US20020037557A1
; GENERAL INFORMATION:
; APPLICANT: Bass, Michael B.
; APPLICANT: Jing, Shuguan
; TITLE OF INVENTION: Fibroblast Growth Factor-like Molecules and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 01-006-A1
; CURRENT APPLICATION NUMBER: US/09/805,805
; CURRENT FILING DATE: 2001-03-13
; PRIOR APPLICATION NUMBER: 60/188,786
; PRIOR FILING DATE: 2000-03-13
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 8
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
; US-09-805-805-8

```

Query Match	100.0%	Score 58	DB 9	Length 11
Best Local Similarity	100.0%	Pred. No. 0.018		
Matches 11	Conservative	0	Mismatches 0	Indels 0
			Gaps 0	

Qy 1 YGRKRRQRRR 11  
|||  
Db 1 YGRKRRQRRR 11

## RESULT 6

US-09-821-821-24  
; Sequence 24, Application US/09821821  
; Patent No. US20020064823A1  
; GENERAL INFORMATION:  
; APPLICANT: Welcher, Andrew A.  
; APPLICANT: Calzone, Frank J.  
; TITLE OF INVENTION: CD20/IgE-Receptor Like Molecules and Uses Thereof  
; FILE REFERENCE: 01017/36938A  
; CURRENT APPLICATION NUMBER: US/09/821-821

; CURRENT FILING DATE: 2001-03-29  
; PRIOR APPLICATION NUMBER: US 09/723,258  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: US 60/193,728  
; PRIOR FILING DATE: 2000-03-30  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 24  
; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Peptide  
US-09-821-821-24

Query Match 100.0%; Score 58; DB 9; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.018;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 YGRKKRQRRR 11  
| | | | | | | | | | |  
Db 1 YGRKKRQRRR 11

RESULT 7  
US-09-895-943-13  
; Sequence 13, Application US/09895943  
; Patent No. US20020068323A1  
; GENERAL INFORMATION:  
; APPLICANT: Sarris, Chris  
; TITLE OF INVENTION: Thymic Stromal Lymphopoietin Receptor Molecules and  
; TITLE OF INVENTION: Uses Thereof  
; FILE REFERENCE: 00-514-C  
; CURRENT APPLICATION NUMBER: US/09/895,943  
; CURRENT FILING DATE: 2001-06-28  
; PRIOR APPLICATION NUMBER: 60/214,866  
; PRIOR FILING DATE: 2000-06-28  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 13  
; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: Human immunodeficiency virus type 1  
US-09-895-943-13

Query Match 100.0%; Score 58; DB 9; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.018;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 YGRKKRQRRR 11  
| | | | | | | | | | |  
Db 1 YGRKKRQRRR 11

RESULT 8  
US-09-825-414-91  
; Sequence 91, Application US/09825414  
; Patent No. US20020083489A1  
; GENERAL INFORMATION:  
; APPLICANT: Collmer, Alan  
; APPLICANT: Alfano, James R.  
; APPLICANT: Charkowski, Amy O.  
; TITLE OF INVENTION: DNA MOLECULES AND POLYPEPTIDES OF PSEUDOMONAS SYRINGAE  
; TITLE OF INVENTION: HRP PATHOGENICITY ISLAND AND THEIR USES  
; FILE REFERENCE: 19603/3243  
; CURRENT APPLICATION NUMBER: US/09/825,414  
; CURRENT FILING DATE: 2001-04-03  
; PRIOR APPLICATION NUMBER: 60/194,160  
; PRIOR FILING DATE: 2000-04-03  
; PRIOR APPLICATION NUMBER: 60/224,604  
; PRIOR FILING DATE: 2000-08-11  
; PRIOR APPLICATION NUMBER: 60/249,548

; PRIOR FILING DATE: 2000-11-17  
; NUMBER OF SEQ ID NOS: 91  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 91  
; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: human  
; OTHER INFORMATION: immunodeficiency virus TAT protein, transduction  
; OTHER INFORMATION: domain  
US-09-825-414-91

Query Match 100.0%; Score 58; DB 9; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.018;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 YGRKKRQRRR 11  
| | | | | | | | | | |  
Db 1 YGRKKRQRRR 11

RESULT 9  
US-09-867-274-23  
; Sequence 23, Application US/09867274  
; Patent No. US20020106650A1  
; GENERAL INFORMATION:  
; APPLICANT: Paszty, Christopher  
; APPLICANT: Gao, Yongming  
; TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof  
; FILE REFERENCE: 01017/37428  
; CURRENT APPLICATION NUMBER: US/09/867,274  
; CURRENT FILING DATE: 2001-05-29  
; PRIOR APPLICATION NUMBER: US 60/208,550  
; PRIOR FILING DATE: 2000-06-01  
; PRIOR APPLICATION NUMBER: US 60/223,542  
; PRIOR FILING DATE: 2000-08-04  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 23  
; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: HIV TAT peptide  
US-09-867-274-23

Query Match 100.0%; Score 58; DB 9; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.018;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 YGRKKRQRRR 11  
| | | | | | | | | | |  
Db 1 YGRKKRQRRR 11

RESULT 10  
US-09-955-866-24  
; Sequence 24, Application US/09955866  
; Patent No. US20020107363A1  
; GENERAL INFORMATION:  
; APPLICANT: Fox, Michael  
; APPLICANT: Sullivan, John K.  
; APPLICANT: Holst, Paige  
; APPLICANT: Yoshinaga, Steven Kiyoshi  
; TITLE OF INVENTION: B7-Like Polypeptides and Uses Thereof  
; FILE REFERENCE: 00,759-A  
; CURRENT APPLICATION NUMBER: US/09/955,866  
; CURRENT FILING DATE: 2001-09-19  
; PRIOR APPLICATION NUMBER: 60/233,867  
; PRIOR FILING DATE: 2000-09-20  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: PatentIn Ver. 2.0

```
; SEQ ID NO 24
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-09-955-866-24

Query Match.      100.0%; Score 58; DB 9; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.018;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 YGRKKRRQRRR 11
Db      1 YGRKKRRQRRR 11
|||||

RESULT 11
US-09-928-175-25
; Sequence 25, Application US/09928175
; Patent No. US20020123618A1
; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher J.
; APPLICANT: Gong, Jiahua
; APPLICANT: Daugherty, Betsy
; APPLICANT: Rogers, No. US20020123618Alma
; TITLE OF INVENTION: Leucine-Rich G Protein Coupled Receptor-8 Molecules and
; TITLE OF INVENTION: Uses Thereof
; FILE REFERENCE: 00-1229
; CURRENT APPLICATION NUMBER: US/09/928,175
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/224,455
; PRIOR FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 25
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-09-928-175-25

Query Match      100.0%; Score 58; DB 9; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.018;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 YGRKKRRQRRR 11
Db      1 YGRKKRRQRRR 11
|||||

RESULT 12
US-09-995-542-13
; Sequence 13, Application US/09995542
; Patent No. US20020127647A1
; GENERAL INFORMATION:
; APPLICANT: Shutter, John
; APPLICANT: Ullas, Laarni
; TITLE OF INVENTION: ATP-Binding Cassette Transporter-Like Molecules and
; TITLE OF INVENTION: Uses Thereof
; FILE REFERENCE: 00-658-A
; CURRENT APPLICATION NUMBER: US/09/995,542
; CURRENT FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: 60/253,520
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-09-995-542-13

Query Match      100.0%; Score 58; DB 9; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.018;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 YGRKKRRQRRR 11
Db      1 YGRKKRRQRRR 11
|||||

RESULT 13
US-09-927-850-18
; Sequence 18, Application US/09927850
; Patent No. US20020137137A1
; GENERAL INFORMATION:
; APPLICANT: Welcher, Andrew
; APPLICANT: Wen, Duanzhi
; APPLICANT: Kelly, Michael
; TITLE OF INVENTION: Interferon-Like Molecules and Uses Thereof
; FILE REFERENCE: 99,372-F
; CURRENT APPLICATION NUMBER: US/09/927,850
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 09/724,860
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 60/169,720
; PRIOR FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-09-927-850-18
```

```
Query Match      100.0%; Score 58; DB 9; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.018;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 YGRKKRRQRRR 11
Db      1 YGRKKRRQRRR 11
|||||
```

```
RESULT 14
US-09-984-183-4
; Sequence 4, Application US/09984183
; Patent No. US20020142983A1
; GENERAL INFORMATION:
; APPLICANT: AGRAWAL, BABITA
; APPLICANT: LONGENECKER, MICHAEL B.
; TITLE OF INVENTION: MUC-1 ANTAGONISTS AND METHODS OF TREATING IMMUNE
; TITLE OF INVENTION: DISORDERS
; FILE REFERENCE: 042881/0130
; CURRENT APPLICATION NUMBER: US/09/984,183
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 09/457,354
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/111,973
; PRIOR FILING DATE: 1998-12-11
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus
US-09-984-183-4
```

```
Query Match      100.0%; Score 58; DB 9; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.018;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 YGRKKRRQRRR 11
Db      1 YGRKKRRQRRR 11
|||||
```

RESULT 15



```

US-09-949-196-15
; Sequence 15, Application US/09949196
; Patent No. US20020147145A1
; GENERAL INFORMATION:
; APPLICANT: Zealand Pharmaceuticals A/S
; TITLE OF INVENTION: MATERIALS AND METHODS RELATING TO THE DEGRADATION OF Cdc25A IN RE
; TITLE OF INVENTION: TO DNA DAMAGE
; FILE REFERENCE: 55888 (45487)
; CURRENT APPLICATION NUMBER: US/09/949,196
; CURRENT FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic peptide sequence
US-09-949-196-15

Query Match      100.0%; Score 58; DB 9; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.018;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 YGRKKRRQRRR 11
        |||||
Db      1 YGRKKRRQRRR 11

```

Search completed: March 31, 2005, 00:34:56  
Job time : 7.28846 secs

1000

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2005, 00:23:16 ; Search time 2.2 Seconds  
(without alignments)  
373.245 Million cell updates/sec

Title: US-10-049-822A-12

Perfect score: 58

Sequence: 1 YGRKKRQRRR 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*

- 1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/iaa/PTJUS\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	58	100.0	11	2	US-08-706-741B-54
2	58	100.0	11	3	US-08-924-635A-54
3	58	100.0	11	3	US-09-208-966-2
4	58	100.0	11	3	US-09-296-089-37
5	58	100.0	11	4	US-09-837-863-2
6	58	100.0	11	4	US-09-660-742-1
7	58	100.0	11	4	US-09-434-345-2
8	58	100.0	11	4	US-09-632-287A-22
9	58	100.0	11	4	US-09-632-277A-4
10	58	100.0	11	4	US-09-612-033B-15
11	58	100.0	11	4	US-09-780-070-37
12	58	100.0	11	4	US-09-775-052A-2
13	58	100.0	11	4	US-09-911-842A-6
14	58	100.0	11	4	US-09-997-465B-2
15	58	100.0	11	4	US-10-083-889-17
16	58	100.0	11	4	US-09-551-976-37
17	58	100.0	11	4	US-09-265-107-75
18	58	100.0	11	4	US-09-724-136A-16
19	58	100.0	11	4	US-09-545-433-14
20	58	100.0	11	4	US-10-031-505-12
21	58	100.0	11	4	US-10-144-549-6
22	58	100.0	12	1	US-08-450-257-47
23	58	100.0	12	1	US-08-450-246-47
24	58	100.0	12	1	US-08-450-098-47
25	58	100.0	12	1	US-08-451-233-47
26	58	100.0	12	1	US-08-450-236-47
27	58	100.0	12	2	US-08-733-505A-50

28 58 100.0 12 3 US-08-235-403-47 Sequence 47, Appl  
29 58 100.0 12 4 US-09-254-126D-46 Sequence 46, Appl  
30 58 100.0 12 4 US-09-656-121-14 Sequence 14, Appl  
31 58 100.0 12 4 US-09-656-121-15 Sequence 15, Appl  
32 58 100.0 14 5 PCT-US95-06077-4 Sequence 4, Appl  
33 58 100.0 14 5 PCT-US95-06077-10 Sequence 10, Appl  
34 58 100.0 14 5 PCT-US95-06077-11 Sequence 11, Appl  
35 58 100.0 14 5 PCT-US95-06077-12 Sequence 12, Appl  
36 58 100.0 15 1 US-08-450-257-5 Sequence 5, Appl  
37 58 100.0 15 1 US-08-450-257-6 Sequence 5, Appl  
38 58 100.0 15 1 US-08-450-246-5 Sequence 5, Appl  
39 58 100.0 15 1 US-08-450-246-6 Sequence 6, Appl  
40 58 100.0 15 1 US-08-450-098-5 Sequence 5, Appl  
41 58 100.0 15 1 US-08-450-098-6 Sequence 6, Appl  
42 58 100.0 15 1 US-08-451-233-5 Sequence 5, Appl  
43 58 100.0 15 1 US-08-451-233-6 Sequence 6, Appl  
44 58 100.0 15 1 US-08-450-236-5 Sequence 5, Appl  
45 58 100.0 15 1 US-08-450-236-6 Sequence 6, Appl

#### ALIGNMENTS

RESULT 1  
US-08-706-741B-54  
; Sequence 54, Application US/08706741B  
; Patent No. 5955593  
; GENERAL INFORMATION:  
; APPLICANT: KORSMEYER, STANLEY J.  
; TITLE OF INVENTION: BH3 INTERACTING DOMAIN DEATH AGONIST  
; NUMBER OF SEQUENCES: 88  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: HOWELL & HAFERKAMP, L.C.  
; STREET: 7733 FORSYTH BLVD., SUITE 1400  
; CITY: ST. LOUIS  
; STATE: MISSOURI  
; COUNTRY: USA  
; ZIP: 63146  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08706741B  
; FILING DATE: 09-SEP-1996  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: HOLLAND, DONALD R.  
; REGISTRATION NUMBER: 35,197  
; REFERENCE/DOCKET NUMBER: 965017  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (314) 727-5188  
; TELEFAX: (314) 727-6092  
; INFORMATION FOR SEQ ID NO: 54:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-706-741B-54

Query Match 100.0%; Score 58; DB 2; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0029;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 YGRKKRQRRR 11  
| | | | | | | | | | | |

Db 1 YGRKKRQRRR 11  
| | | | | | | | | | | |

RESULT 2

US-08-924-695A-54  
; Sequence 54, Application US/08924695A  
; Patent No. 598583  
; GENERAL INFORMATION:  
; APPLICANT: KORSMEYER, STANLEY J.  
; TITLE OF INVENTION: BH3 INTERACTING DOMAIN DEATH AGONIST  
; NUMBER OF SEQUENCES: 88  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: HOWELL & HAFERKAMP, L.C.,  
; STREET: 7733 FORSYTH BLVD., SUITE 1400  
; CITY: ST. LOUIS  
; STATE: MISSOURI  
; COUNTRY: USA  
; ZIP: 63105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/924,695A  
; FILING DATE: 09-SEP-1997  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: HOLLAND, DONALD R.  
; REGISTRATION NUMBER: 35,197  
; REFERENCE/DOCKET NUMBER: 971798  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (314) 727-5188  
; TELEFAX: (314) 727-6092  
; INFORMATION FOR SEQ ID NO: 54:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-924-695A-54

Query Match 100.0%; Score 58; DB 2; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0029;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YGRKKRRQRRR 11  
| | | | | | | | | |  
Db 1 YGRKKRRQRRR 11

RESULT 3  
US-09-208-966-2  
; Sequence 2, Application US/09208966  
; Patent No. 6221355  
; GENERAL INFORMATION:  
; APPLICANT: Dowdy, Steven F.  
; TITLE OF INVENTION: ANTI-PATHOGEN SYSTEM AND METHODS OF USE THEREOF  
; FILE REFERENCE: 48881/1742  
; CURRENT APPLICATION NUMBER: US/09/208,966  
; CURRENT FILING DATE: 1998-12-10  
; EARLIER APPLICATION NUMBER: 60/082,402  
; EARLIER FILING DATE: 1998-04-20  
; EARLIER APPLICATION NUMBER: 60/069,012  
; EARLIER FILING DATE: 1997-12-10  
; NUMBER OF SEQ ID NOS: 57  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: human  
US-09-208-966-2

Query Match 100.0%; Score 58; DB 3; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0029;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YGRKKRRQRRR 11  
| | | | | | | | | |  
Db 1 YGRKKRRQRRR 11

RESULT 4  
US-09-296-089-37  
; Sequence 37, Application US/09296089  
; Patent No. 6303576  
; GENERAL INFORMATION:  
; APPLICANT: Blaschuk, Orest W.  
; APPLICANT: Byers, Stephen  
; APPLICANT: Gour, Barbara J.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING  
; TITLE OF INVENTION: BETA-CATENIN MEDIATED GENE EXPRESSION  
; FILE REFERENCE: 100086.411  
; CURRENT APPLICATION NUMBER: US/09/296,089  
; CURRENT FILING DATE: 1999-04-21  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 37  
; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: Human immunodeficiency virus  
US-09-296-089-37

Query Match 100.0%; Score 58; DB 3; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0029;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YGRKKRRQRRR 11  
| | | | | | | | | |  
Db 1 YGRKKRRQRRR 11

RESULT 5  
US-09-837-863-2  
; Sequence 2, Application US/09837863  
; Patent No. 6468754  
; GENERAL INFORMATION:  
; APPLICANT: Greene, Amy  
; APPLICANT: Zhou, Hua  
; APPLICANT: Thode, Silke  
; APPLICANT: Jarnigan, Kurt  
; TITLE OF INVENTION: Vector and Method for Targeted Replacement and Disruption  
; TITLE OF INVENTION: of an Integrated DNA Sequence  
; FILE REFERENCE: 025.1US  
; CURRENT APPLICATION NUMBER: US/09/837,863  
; CURRENT FILING DATE: 2001-04-17  
; PRIOR APPLICATION NUMBER: US 60/198,498  
; PRIOR FILING DATE: 2000-04-18  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 2  
; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: Human immunodeficiency virus  
US-09-837-863-2

Query Match 100.0%; Score 58; DB 4; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0029;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YGRKKRRQRRR 11  
| | | | | | | | | |  
Db 1 YGRKKRRQRRR 11

RESULT 6  
US-09-660-742-1  
; Sequence 1, Application US/09660742  
; Patent No. 6485977

; GENERAL INFORMATION:  
; APPLICANT: Collmer, Alan  
; APPLICANT: Beer, Steven V.  
; TITLE OF INVENTION: RECOMBINANT CONSTRUCTS AND TECHNIQUES FOR DELIVERING TO  
; TITLE OF INVENTION: EUCARYOTIC CELLS BACTERIAL PROTEINS THAT ARE SECRETED  
; TITLE OF INVENTION: VIA TYPE III SECRETION SYSTEMS  
; FILE REFERENCE: 19603/4021  
; CURRENT APPLICATION NUMBER: US/09/660,742  
; CURRENT FILING DATE: 2000-09-13  
; PRIOR APPLICATION NUMBER: 60/153,507  
; PRIOR FILING DATE: 1999-09-13  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-660-742-1

Query Match 100.0%; Score 58; DB 4; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0029;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YGRKKRQRRR 11  
| | | | | | | | | |  
Db 1 YGRKKRQRRR 11

RESULT 7  
US-09-434-345-2  
; Sequence 2, Application US/09434345  
; Patent No. 6511676  
; GENERAL INFORMATION:  
; APPLICANT: Boulikas, Teni  
; TITLE OF INVENTION: THERAPY FOR HUMAN CANCERS USING  
; TITLE OF INVENTION: CISPLATIN AND OTHER DRUGS OR GENES ENCAPSULATED INTO  
; TITLE OF INVENTION: LIPOSOMES  
; FILE REFERENCE: TB 2001.00  
; CURRENT APPLICATION NUMBER: US/09/434,345  
; CURRENT FILING DATE: 1999-11-05  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: Human Immunodeficiency Virus  
US-09-434-345-2

Query Match 100.0%; Score 58; DB 4; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0029;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YGRKKRQRRR 11  
| | | | | | | | | |  
Db 1 YGRKKRQRRR 11

RESULT 8  
US-09-632-287A-22  
; Sequence 22, Application US/09632287A  
; Patent No. 6521422  
; GENERAL INFORMATION:  
; APPLICANT: Hsu, Hailing  
; APPLICANT: Wooden, Scott K  
; APPLICANT: Boyle, William J  
; TITLE OF INVENTION: Fhm, A No. 6521422el Member of the TNF Ligand Supergene Family  
; FILE REFERENCE: 01017/35550A  
; CURRENT APPLICATION NUMBER: US/09/632,287A  
; CURRENT FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: US 60/147,294  
; PRIOR FILING DATE: 1999-08-04  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 22  
; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Peptide from the HIV TAT prot  
US-09-632-287A-22

Query Match 100.0%; Score 58; DB 4; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0029;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YGRKKRQRRR 11  
| | | | | | | | | |  
Db 1 YGRKKRQRRR 11

RESULT 9  
US-09-632-277A-4  
; Sequence 4, Application US/09632277A  
; Patent No. 6599716  
; GENERAL INFORMATION:  
; APPLICANT: Hsu, Hailing  
; TITLE OF INVENTION: NTR3 A No. 6599716el Member of the TNF-Receptor Supergene Family  
; FILE REFERENCE: 01017/3549A  
; CURRENT APPLICATION NUMBER: US/09/632,277A  
; CURRENT FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: US 60/147,297  
; PRIOR FILING DATE: 1999-08-04  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: HIV TAT peptide  
US-09-632-277A-4

Query Match 100.0%; Score 58; DB 4; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0029;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YGRKKRQRRR 11  
| | | | | | | | | |  
Db 1 YGRKKRQRRR 11

RESULT 10  
US-09-612-033B-15  
; Sequence 15, Application US/09612033B  
; Patent No. 6627199  
; GENERAL INFORMATION:  
; APPLICANT: Saris, Chris  
; TITLE OF INVENTION: Isolation, Identification, and Characterization of  
; TITLE OF INVENTION: tmst2, a No. 6627199el Member of the TNF-Receptor Superfamily  
; FILE REFERENCE: 01017/35434A  
; CURRENT APPLICATION NUMBER: US/09/612,033B  
; CURRENT FILING DATE: 2000-07-07  
; PRIOR APPLICATION NUMBER: US 60/143,063  
; PRIOR FILING DATE: 1999-07-09  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 15  
; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: peptide  
US-09-612-033B-15

Query Match 100.0%; Score 58; DB 4; Length 11;

Best Local Similarity 100.0%; Pred. No. 0.0029;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YGRKKRQRRR 11  
| | | | | | | | | |  
Db 1 YGRKKRQRRR 11

## RESULT 11

US-09-780-070-37  
; Sequence 37, Application US/09780070  
; Patent No. 6632616  
; GENERAL INFORMATION:  
; APPLICANT: Burke, James  
; APPLICANT: Strittmatter, Warren  
; APPLICANT: Nagai, Yoshitaka  
; TITLE OF INVENTION: COMPOUNDS THAT SELECTIVELY BIND TO EXPANDED POLYGLUTAMINE REPEAT  
; AND METHODS OF USE THEREOF  
; FILE REFERENCE: 5405.242  
; CURRENT APPLICATION NUMBER: US/09/780,070  
; CURRENT FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/189,781  
; PRIOR FILING DATE: 2000-03-16  
; NUMBER OF SEQ ID NOS: 40  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 37  
; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: Human immunodeficiency virus  
US-09-780-070-37

Query Match 100.0%; Score 58; DB 4; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0029;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YGRKKRQRRR 11  
| | | | | | | | | |  
Db 1 YGRKKRQRRR 11

## RESULT 12

US-09-775-052A-2  
; Sequence 2, Application US/09775052A  
; Patent No. 6645501  
; GENERAL INFORMATION:  
; APPLICANT: Dowdy, Steven F.  
; TITLE OF INVENTION: ANTI-PATHOGEN SYSTEM AND METHODS OF USE THEREOF  
; FILE REFERENCE: 48881/1742  
; CURRENT APPLICATION NUMBER: US/09/775,052A  
; CURRENT FILING DATE: 2001-12-05  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/208,966  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-10  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/069,012  
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-12-10  
; NUMBER OF SEQ ID NOS: 57  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: human  
US-09-775-052A-2

Query Match 100.0%; Score 58; DB 4; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0029;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YGRKKRQRRR 11  
| | | | | | | | | |  
Db 1 YGRKKRQRRR 11

## RESULT 13

US-09-911-842A-6

; Sequence 6, Application US/09911842A  
; Patent No. 6656707  
; GENERAL INFORMATION:  
; APPLICANT: Amgen Inc.  
; TITLE OF INVENTION: C3B/C4B COMPLEMENT RECEPTOR-LIKE MOLECULES AND USES THEREOF  
; FILE REFERENCE: 01017/37592  
; CURRENT APPLICATION NUMBER: US/09/911,842A  
; CURRENT FILING DATE: 2001-07-24  
; PRIOR APPLICATION NUMBER: US 60/222,438  
; PRIOR FILING DATE: 2000-08-01  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 6  
; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-911-842A-6

Query Match 100.0%; Score 58; DB 4; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0029;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YGRKKRQRRR 11  
| | | | | | | | | |  
Db 1 YGRKKRQRRR 11

## RESULT 14

US-09-997-465B-2  
; Sequence 2, Application US/09997465B  
; Patent No. 6673574  
; GENERAL INFORMATION:  
; APPLICANT: Stern, William  
; APPLICANT: Mehta, No. 6673574er M.  
; APPLICANT: Ray, Martha V.L.  
; TITLE OF INVENTION: IMPROVED ORAL DELIVERY OF PEPTIDES USING ENZYME-CLEAVABLE MEMBRAN  
; TITLE OF INVENTION: TRANSLOCATORS  
; FILE REFERENCE: P/546-247  
; CURRENT APPLICATION NUMBER: US/09/997,465B  
; CURRENT FILING DATE: 2001-11-29  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: Human immunodeficiency virus type 1  
US-09-997-465B-2

Query Match 100.0%; Score 58; DB 4; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0029;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YGRKKRQRRR 11  
| | | | | | | | | |  
Db 1 YGRKKRQRRR 11

## RESULT 15

US-10-083-889-17  
; Sequence 17, Application US/10083889  
; Patent No. 6673894  
; GENERAL INFORMATION:  
; APPLICANT: Zahner, Joseph E.  
; TITLE OF INVENTION: Inhibitor of cell proliferation and methods of use thereof.  
; FILE REFERENCE: 16850-7331  
; CURRENT APPLICATION NUMBER: US/10/083,889  
; CURRENT FILING DATE: 2002-02-27  
; PRIOR APPLICATION NUMBER: US 60/271,798  
; PRIOR FILING DATE: 2001-02-27  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: Microsoft Word 97  
; SEQ ID NO 17  
; LENGTH: 11

```
; TYPE: PRT
; ORGANISM: HIV-1
US-10-083-889-17

Query Match      100.0%; Score 58; DB 4; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.0029;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 YGRKKRQRRR 11
      |||||
Db      1 YGRKKRQRRR 11

Search completed: March 31, 2005, 00:32:39
Job time : 4.2 secs
```

Figure 1



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2005, 00:17:06 ; Search time 7.7 Seconds  
(without alignments)  
552.515 Million cell updates/sec

Title: US-10-049-822A-12

Perfect score: 58

Sequence: 1 YGRKRRQRRR 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

1: Geneseqp1980s:\*

2: Geneseqp1990s:\*

3: Geneseqp2000s:\*

4: Geneseqp2001s:\*

5: Geneseqp2002s:\*

6: Geneseqp2003as:\*

7: Geneseqp2003bs:\*

8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	58	100.0	11	2	AAW50263 HIV-1 tat
2	58	100.0	11	2	AAW50263
3	58	100.0	11	2	AAW50263
4	58	100.0	11	2	AAW50263
5	58	100.0	11	3	AAW50263
6	58	100.0	11	3	AAW50263
7	58	100.0	11	3	AAW50263
8	58	100.0	11	3	AAW50263
9	58	100.0	11	3	AAW50263
10	58	100.0	11	3	AAW50263
11	58	100.0	11	3	AAW50263
12	58	100.0	11	4	AAW50263
13	58	100.0	11	4	AAW50263
14	58	100.0	11	4	AAW50263
15	58	100.0	11	4	AAW50263
16	58	100.0	11	4	AAW50263
17	58	100.0	11	4	AAW50263
18	58	100.0	11	4	AAW50263
19	58	100.0	11	4	AAW50263
20	58	100.0	11	4	AAW50263
21	58	100.0	11	4	AAW50263
22	58	100.0	11	4	AAW50263
23	58	100.0	11	4	AAW50263
24	58	100.0	11	4	AAW50263
25	58	100.0	11	4	AAW50263

26	58	100.0	11	4	AAE03730	Protein t
27	58	100.0	11	4	AAW50221	AAW50221 HIV-1 tat
28	58	100.0	11	4	AAU09932	AAU09932 Human imm
29	58	100.0	11	4	AAE12891	AAE12891 Human imm
30	58	100.0	11	4	AAE13064	AAE13064 Protein t
31	58	100.0	11	4	AAE13064	AAE13064 Protein t
32	58	100.0	11	4	AAE13064	AAE13064 Protein t
33	58	100.0	11	4	AAE13064	AAE13064 Protein t
34	58	100.0	11	4	AAE13064	AAE13064 Protein t
35	58	100.0	11	4	AAE13064	AAE13064 Protein t
36	58	100.0	11	4	AAE13064	AAE13064 Protein t
37	58	100.0	11	4	AAE13064	AAE13064 Protein t
38	58	100.0	11	4	AAE13064	AAE13064 Protein t
39	58	100.0	11	4	AAE13064	AAE13064 Protein t
40	58	100.0	11	4	AAE13064	AAE13064 Protein t
41	58	100.0	11	4	AAE13064	AAE13064 Protein t
42	58	100.0	11	4	AAE13064	AAE13064 Protein t
43	58	100.0	11	4	AAE13064	AAE13064 Protein t
44	58	100.0	11	4	AAE13064	AAE13064 Protein t
45	58	100.0	11	4	AAE13064	AAE13064 Protein t

#### ALIGNMENTS

RESULT 1

AAW50263

ID AAW50263 standard; protein; 11 AA.

XX AAW50263;

AC AAW50263;

XX 17-OCT-2003 (revised)

DT 20-JUL-1998 (first entry)

XX HIV-1 tat protein.

XX Mouse; BH3 interacting domain death agonist; BID; BCL-2 family;

XX apoptosis; regulation; cell death; inflammation; cancer; arthritis;

XX autoimmune disease; viral infection; lymphoproliferative.

XX Human immunodeficiency virus 1.

XX WO9809980-A1.

XX 12-MAR-1998.

XX 09-SEP-1997; 97WO-US015872.

XX 09-SEP-1996; 96US-00706741.

XX (UNIW ) UNIV WASHINGTON.

XX Korsmeyer SJ;

XX WPI; 1998-193546/17.

XX BH3 interacting domain death agonist polypeptide - used for treating

XX decreased apoptotic conditions resulting from inflammation etc.

XX Example 8; Page 85; 118pp; English.

XX The present sequence represents the HIV-1 tat protein which is used in an

XX example of the present invention which describes a BH3 interacting domain

XX death agonist (BID) truncated protein. The BID protein, the DNA encoding

XX it or antisense sequences can be used for preventing or treating a

XX decreased apoptotic state of a cell. The decreased apoptotic state that

XX is treated results from a disease such as cancer, viral infections,

XX lymphoproliferative conditions, arthritis, inflammation and autoimmune

XX diseases. Antibodies against the BID protein can be used for detecting a

XX BID polypeptide in a cell or population of cell. The nucleic acid

XX sequence and the BID protein can also be used for treating

XX immunodeficiency disease (including AIDS), senescence, neurodegenerative

XX disease, ischaemic and reperfusion cell death, infertility and wound-

CC healing. Primers derived from the nucleic acid encoding the BID protein  
CC can be used for detecting/quantitating the protein and for detecting  
CC alterations in the nucleic acid encoding the BID protein. (Updated on 17-  
CC OCT-2003 to standardise OS field)

XX  
XX  
SQ Sequence 11 AA;

Query Match 100.0%; Score 58; DB 2; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.013;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YGRKKRQRRR 11  
| | | | | | | | | |  
Db 1 YGRKKRQRRR 11

RESULT 2  
AA05415  
ID AAY05415 standard; peptide; 11 AA.

XX  
XX  
AC AAY05415;

DT 02-JUL-1999 (first entry)

DE Tat peptide.

XX BH3 domain; cell death agonist; bcl homology domain; BCL-2 family;  
KW apoptosis promoter; cancer cell; virus infected cell; inflammation;  
KW autoantibody producing cell; cancer; lymphoproliferative condition;  
KW arthritis; autoimmune disease; therapy.

XX Unidentified.

PN WO9916787-A1.

XX  
XX  
PD 08-APR-1999.

XX 22-SEP-1998; 98WO-US019765.

PR 26-SEP-1997; 97US-0060133P.

PR 07-OCT-1997; 97US-00946039.

XX (UNIW ) UNIV WASHINGTON.

XX Koresmeyer SJ;

XX WPI; 1999-255058/21.

XX Bcl homology domain 3 polypeptide.

XX Claim 7; Page 67; 104pp; English.

XX This sequence represents a Tat peptide. The invention relates to a bcl  
CC homology domain 3 (BH3 domain), derived from a proapoptotic member of the  
CC BCL-2 family. The BH3 polypeptide can be used in a method for promoting  
CC apoptosis in a target cell, especially where the cell is a cancer cell a  
CC virus infected cell or an autoantibody producing cell. The BH3  
CC polypeptide can be used in therapeutic compositions for treating disease  
CC including cancer, other lymphoproliferative conditions, arthritis,  
CC inflammation, and autoimmune diseases, which may result from the down  
CC regulation of cell death regulation

XX Sequence 11 AA;

Query Match 100.0%; Score 58; DB 2; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.013;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YGRKKRQRRR 11  
| | | | | | | | | |  
Db 1 YGRKKRQRRR 11

RESULT 3  
AAY25075

ID AAY25075 standard; peptide; 11 AA.

XX  
XX  
AC AAY25075;

DT 24-AUG-1999 (first entry)

DE Tat transduction domain peptide motif.

XX Anti-pathogen; fusion protein; protein transduction domain; PTD; AZT;  
KW cytotoxic domain; suppressor; infection; medicament; ddi; ddc; d4T; 3TC;  
KW FTC; DAPD; 1592U89; CS92; acyclovir; ganciclovir; peniclovir; interferon;  
KW apoptosis; virus; HIV; cytomegalovirus; CMV; herpes simplex virus; HSV-1;  
KW hepatitis virus; Kaposi's sarcoma-associated herpes virus; KSHV; dial;  
KW herpes virus; yellow fever virus; flavivirus; rhinovirus; plasmodial;  
KW transduction efficiency; cytotoxin.

XX Unidentified.

XX WO9929721-A1.

XX 17-JUN-1999.

XX 10-DEC-1998; 98WO-US026358.

XX 10-DEC-1997; 97US-0069012P.

XX 20-APR-1998; 98US-0082402P.

XX (UNIW ) UNIV WASHINGTON.

XX Dowdy SF;

XX WPI; 1999-394958/33.

XX New anti-pathogen systems, particularly for virus and plasmodium  
XX infections.

XX Claim 65; Page 34; 123pp; English.

XX This invention describes a novel anti-pathogen system (APS) comprising a  
CC fusion protein constructed from a covalently linked protein transduction  
CC domain (PTD) and a cytotoxic domain. The APS can be used for suppressing  
CC a pathogen infection in a mammal. The method may further comprise  
CC administering a medicament e.g. AZT, ddi, ddc, d4T, 3TC, FTC, DAPD,  
CC 1592U89, CS92, acyclovir, ganciclovir, peniclovir or an interferon. The  
CC APS can also be administered to a mammal in the presence of a pathogen to  
CC induce apoptosis in a predetermined population of cells. The products can  
CC be used for treating mammals suffering from or susceptible to a viral  
CC infection or a disease associated with a virus, e.g. HIV, cytomegalovirus  
CC (CMV), herpes simplex virus, e.g. type 1 (HSV-1) hepatitis virus, type C  
CC (HCV), Kaposi's sarcoma-associated herpes virus (KSHV or human herpes  
CC virus 8), yellow fever virus, flavivirus or rhinovirus, or suffering from  
CC or susceptible to plasmodial infection or a disease associated with a  
CC plasmodial infection, e.g. P. falciparum, P. vivax, P. ovale, or P.  
CC malariae. The APS exhibits high transduction efficiency and specifically  
CC kills or injures cells infected by one or more pathogens. Formation of  
CC the cytotoxin is minimized or eliminated in uninfected cells and in  
CC infected cells that keep the pathogen inactive. The APS can be  
CC specifically tailored to kill or injure cells infected by one or more  
CC pathogen strains. This sequence represents a Tat transduction domain  
CC peptide motif which is used in the method of the invention

XX Sequence 11 AA;

Query Match 100.0%; Score 58; DB 2; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.013;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YGRKKRQRRR 11  
| | | | | | | | | |  
Db 1 YGRKKRQRRR 11

RESULT 4  
 ID AAB27088 standard; peptide; 11 AA.  
 AC AAB27088;  
 DT 15-FEB-2001 (first entry)  
 DE Beta-catenin derived internalisation moiety SEQ ID NO: 75.  
 KW Beta-catenin; cadherin-mediated intercellular adhesion;  
 KW cell differentiation; modulating agent; hair loss; skin exfoliation;  
 KW internalisation moiety; flanking sequence; transcription; hearing loss.  
 OS Human immunodeficiency virus.  
 FN WO200053632-A1.  
 PD 14-SEP-2000.  
 PF 07-MAR-2000; 2000WO-CA000222.  
 PR 09-MAR-1999; 99US-00265107.  
 PA (UYMC-) UNIV MCGILL.  
 PI Blaschuk OW, Gour BJ;  
 DR WPI; 2000-594308/56.  
 KW Stimulating beta-catenin mediated gene expression, cellular  
 KW differentiation and hair growth, involves contacting cells with  
 KW modulating agent capable of inhibiting interaction between alpha and beta  
 KW catenin.  
 PS Disclosure; Page 19; 77pp; English.  
 CC The present invention is concerned with methods of modulating the amount  
 CC of free beta-catenin in the cell, and methods of stimulating the  
 CC expression of genes involved in cellular differentiation, the  
 CC transcription of which is under the control of beta-catenin. The peptides  
 CC given in AAB27053-B27088, AAB27284-B27300 and AAB27330-B27351 can be used  
 CC as modulating agents which interrupt the interaction between alpha and  
 CC beta catenin, causing increased levels of the latter and stimulating the  
 CC activation of beta-catenin mediated transcription. This can be used to  
 CC stimulate cell differentiation, which can then be used to promote hair  
 CC growth and skin exfoliation. This latter is particularly useful in the  
 CC improvement of photodamaged skin and to minimise wrinkles. The modulating  
 CC peptide can also be used to reduce hearing loss resulting from inner ear  
 CC disorders such as hyperacusis and tinnitus

Query Match 100.0%; Score 58; DB 3; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.013;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 YGRKKRRQRRR 11  
 DB 1 YGRKKRRQRRR 11

RESULT 5  
 ID AAB09907 standard; peptide; 11 AA.  
 AC AAB09907;  
 DT 06-NOV-2000 (first entry)  
 DE HIV tat protein transduction domain.  
 KW HIV tat protein transduction domain.

KW Targeting signal; MUC-1; immunosuppression; autoimmune disorder;  
 KW immune disorder; inflammatory disorder.  
 OS Human immunodeficiency virus.  
 FN WO200034468-A2.  
 PD 15-JUN-2000.  
 PF 09-DEC-1999; 99WO-US029016.  
 PR 11-DEC-1998; 98US-0111973P.  
 PA (BIOM-) BIOMIRA INC.  
 PI Agrawal B, Longenecker BM;  
 DR WPI; 2000-423418/36.  
 KW Use of agent capable of intracellularly inhibiting mucin MUC-1 for  
 KW inducing T-cell-based immunosuppression and for treating autoimmune  
 KW disorders, transplant rejection and inflammatory disorders.  
 PS Disclosure; Page 15; 51pp; English.  
 CC The present sequence is the tat protein transduction domain from HIV,  
 CC which can be used as a targeting signal. It can be used to internalise  
 CC sequences, such as MUC-1 antagonists, within the cell. MUC-1 is a  
 CC immunosuppressor, and antagonists act to reduce overactive immune  
 CC responses. Thus, the peptide can be used to treat inflammatory disorders  
 CC such as rheumatoid arthritis, psoriasis, allergic contact dermatitis and  
 CC ankylosing spondylitis, autoimmune disorders including myasthenia gravis,  
 CC systemic lupus erythematosus, polyarteritis nodosa, Goodpasture's  
 CC syndrome, isopathic thrombocytopenic purpura, autoimmune haemolytic  
 CC anaemia, Graves' disease, rheumatic fever, pernicious anaemia, insulin-  
 CC resistant diabetes mellitus, bullous pemphigoid, pemphigus vulgaris,  
 CC viral myocarditis, autoimmune thyroiditis, male infertility, sarcoidosis,  
 CC allergic encephalomyelitis, multiple sclerosis, Sjorgens disease,  
 CC Reiter's disease, Celiac disease, sympathetic ophthalmia and primary  
 CC biliary cirrhosis, immune disorders, graft versus host disease and  
 CC transplant rejection

Query Match 100.0%; Score 58; DB 3; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.013;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 YGRKKRRQRRR 11  
 DB 1 YGRKKRRQRRR 11

RESULT 6  
 ID AAY93542 standard; peptide; 11 AA.  
 AC AAY93542;  
 DT 25-SEP-2000 (first entry)  
 DE Amino acid sequence of a synthetic protein transduction domain.  
 KW Protein transduction system; protein transduction domain;  
 KW cytotoxic domain; pathogen infection; retroviral infection;  
 KW plasmidial infection; cancer; prostate cancer.  
 OS Synthetic.  
 FN WO200034308-A2.  
 PD 15-JUN-2000.

PF 10-DEC-1999; 99WO-US029289.  
PR 10-DEC-1998; 98US-0111701P.  
XX (UNIW ) UNIV WASHINGTON.  
PA Dowdy SF;  
XX WPI; 2000-431269/37.  
XX Protein transduction system for treating cancer and pathogenic infections  
PT has a fusion protein comprising a protein transduction domain covalently  
PT linked to a cytotoxic domain.  
XX Example 10; Page 71; 127pp; English.  
XX AAY93542-51 represent synthetic protein transduction domains, which are  
CC used in the protein transduction system of the invention. The  
CC specification describes a protein transduction system, which comprises a  
CC fusion protein. This fusion protein has a covalently linked protein  
CC transduction domain and cytotoxic domain. The system is useful for  
CC treating pathogen infection in mammals, infections such as those caused  
CC by CMV, HSV-1, HCV, KSHV, yellow fever virus, flavivirus or rhinovirus,  
CC retroviral infections such as HIV-1, HIV-2, HTLV-3 and/or LAV, plasmodial  
CC infections associated with P.faciaparum, P.vivax, P.ovale, P.malariae. It  
CC is also useful for treating cancer, especially prostate cancer  
XX Sequence 11 AA;  
SQ

Query Match 100.0%; Score 58; DB 3; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.013;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 YGRKKRQRRR 11  
Db 1 YGRKKRQRRR 11

RESULT 7  
AAY71015  
ID AAY71015 standard; peptide; 11 AA.  
XX AAY71015;  
XX 29-AUG-2000 (first entry)  
DE Human immunodeficiency virus TAT peptide.  
XX Phospholamban; PLB; sarcoplasmic reticulum Ca 2+ ATPase; SERCA2a; HIV;  
KW TAT protein; penetratin; transport peptide; cardiomyocyte; inhibitor;  
KW cargo peptide; contractilin; cardiac contractility; cardiac; treatment;  
KW cardiac disease; heart failure; myocardial dysfunction.  
XX Human immunodeficiency virus.  
OS  
XX WO200025804-A2.  
PN 11-MAY-2000.  
XX 02-NOV-1999; 99WO-US025692.  
XX 02-NOV-1998; 98US-0106718P.  
PR 27-JUL-1999; 99US-0145883P.  
XX (REGC ) UNIV CALIFORNIA.  
PA Chien K, Dillman W, Minamisawa S, He H, Hoshijima M, Meyer M;  
PI Scott C, Wang Y, Silverman GJ;  
XX WPI; 2000-365393/31.  
XX Treating cardiac diseases, e.g. heart failure or myocardial dysfunction  
PT comprises enhancing cardiac contractility by inhibiting interaction

PT between phospholamban and sarcoplasmic reticulum calcium 2+ adenosine  
PT triphosphatase.  
XX Example 5; Page 54; 56pp; English.  
XX The patent discloses a method for the treatment of heart failure, using  
CC small peptide complexes and recombinant proteins, that induces  
CC phospholamban (PLB) deficiency and inhibits the interaction between PLB  
CC and sarcoplasmic reticulum Ca 2+ ATPase (SERCA2a) within cardiomyocytes.  
CC The peptide complex comprises of transport peptide like penetratin and  
CC cargo peptide selected from mutant PLB, native PLB or antibody against  
CC PLB protein (contractilin). Penetratin-PLB peptide functions as a  
CC dominant negative inhibitor of PLB-SERCA2a interaction, enhances cardiac  
CC contractility and reduces blood pressure. This method is useful for the  
CC treatment of cardiac disease e.g. heart failure and myocardial  
CC dysfunction. The present sequence is the amino terminal, 11-amino acid  
CC protein transduction domain of the denatured human immunodeficiency virus  
CC (HIV) TAT protein. It belongs to the penetratin class of peptides, with  
CC translocating properties having the ability to carry hydrophilic  
CC compounds across the plasma membrane. It is used in the construction of  
CC recombinant PLB peptides (e.g., TAT-PLB)  
XX Sequence 11 AA;  
SQ

Query Match 100.0%; Score 58; DB 3; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.013;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 YGRKKRQRRR 11  
Db 1 YGRKKRQRRR 11

RESULT 8  
AAB35698  
ID AAB35698 standard; peptide; 11 AA.  
XX AAB35698;  
XX 12-SEP-2003 (revised)  
DT 16-FEB-2001 (first entry)  
XX Peptide associated with modified beta-catenin expression #31.  
DE Beta-catenin; cell differentiation; hair growth; cancer;  
KW Alzheimer's disease.  
XX Human immunodeficiency virus 1.  
OS  
XX WO200063246-A2.  
PN 26-OCT-2000.  
PD 21-APR-2000; 2000WO-US010753.  
XX 21-APR-1999; 99US-00296089.  
PR (ADHE-) ADHEREX TECHNOLOGIES INC.  
XX Blaschuk OW, Byers S, Gour BJ;  
PI WPI; 2000-679589/66.  
XX Use of modulating agent comprising internalization moiety and a peptide,  
PT for modulating beta-catenin mediated gene transcription and cell  
PT differentiation, for treating cancer, and for inhibiting Alzheimer's  
PT disease.  
XX Claim 6; Page 26; 47pp; English.  
XX The present invention relates to a method for modulating beta-catenin  
CC mediated gene transcription in a cell. The method involves contacting a  
CC cell with a modulating agent comprising an internalization moiety and a

CC peptide comprising a sequence LXXLL or peptide analogue of LXXLL. The  
CC method is useful for modulating beta-catenin mediated gene transcription,  
CC cell differentiation, hair growth, and retinoic acid activity, for  
CC treating cancer, and for inhibiting the development of Alzheimer's  
CC disease. (Updated on 12-SEP-2003 to standardise OS field)

XX Sequence 11 AA;

Query Match 100.0%; Score 58; DB 3; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.013;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YGRKKRQRRR 11  
| | | | | | | | | |  
Db 1 YGRKKRQRRR 11

#### RESULT 9

AAB03961  
ID AAB03961 standard; peptide; 11 AA.

XX AAB03961;

XX 26-FEB-2001 (first entry)

XX Minimal eleven amino acids present in pTAT.

XX Chimeric protein; fusion protein; FLICE like inhibitor protein; FLIP;  
KW Fas; TNF; apoptosis; caspase-8; ligand; T cell; thymocyte;  
KW tumour specific antigen; immune response; therapy; prophylaxis;  
KW diagnosis; HIV; human immunodeficiency syndrome; AIDS;  
KW acquired immune deficiency syndrome.

XX Human immunodeficiency virus.

OS Human immunodeficiency virus.

XX WO200059935-A1.

XX 12-OCT-2000.

XX 05-APR-2000; 2000WO-US009002.

XX 05-APR-1999; 99US-0127867P.

XX 06-APR-1999; 99US-0128021P.

XX (MAYO-) MAYO FOUND MEDICAL EDUCATION & RES.

XX (PAYA/) PAYA C.

XX (ALGE/) ALGECIRAS-SCHMINICH A.

XX Paya C, Algeciras-Schminich A;

XX WPI; 2000-664988/64.

XX Fusion polypeptide useful for inhibiting ligand-induced apoptosis,

XX comprises portion of anti-apoptotic polypeptide linked to a transport

XX group.

XX Disclosure; Page 30; 89pp; English.

XX A chimeric group or fusion peptide which comprises a portion of an anti-  
CC apoptotic polypeptide which inhibits apoptosis of lymphocytes in  
CC combination with a transport group is described. The transport group is  
CC capable of transporting the chimeric group or fusion peptide across the  
CC cell membrane. The anti-apoptotic polypeptide is FLICE-like inhibitor  
CC protein (FLIP) which inhibits Fas and TNF mediated apoptosis by  
CC inhibiting binding of Caspase-8 to the Fas receptor complex, thus  
CC shutting off the downstream Fas signalling pathway. The chimeric group  
CC and fusion peptide are useful for inhibiting ligand-induced apoptosis by  
CC bringing them into contact with T cells. The chimeric group is useful for  
CC expanding T cells in vitro e.g. T cells specific for particular antigens  
CC such as tumour-specific antigen, for enhancing immune response and to  
CC inhibit the apoptosis of chronically activated T cells e.g. activated  
CC CD4<sup>+</sup> T cells in HIV infected patients. The chimeric group is also useful  
CC for therapeutic, prophylactic or diagnosis of intracellular delivery of

CC small molecules and macromolecules such as anti-apoptotic polypeptides  
CC and nucleic acids encoding such polypeptides. Two primers (AAA54297,  
CC AAA54298) were used to amplify the FLIP cDNA for subcloning into the XhoI  
CC -NcoI site of the pHA-TAT vector which contains the N-terminal protein  
CC transduction domain from the human immunodeficiency virus tat protein.  
CC Tat is a preferred transport moiety

XX Sequence 11 AA;

Query Match 100.0%; Score 58; DB 3; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.013;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YGRKKRQRRR 11  
| | | | | | | | | |  
Db 1 YGRKKRQRRR 11

#### RESULT 10

AAB29413  
ID AAB29413 standard; peptide; 11 AA.

XX AAB29413;

XX 09-FEB-2001 (first entry)

XX HIV TAT transduction domain, SEQ ID NO:2.

XX Protein transduction domain; fusion molecule; therapeutic agent;  
KW drug targeting; drug discovery; cell transduction; bioavailability;  
KW vaccine; nervous system disorder; Alzheimer's disease;  
KW Parkinson's disease; Huntington's disease; pre-senile dementia; epilepsy;  
KW seizure; compulsive behaviour; meningitis; encephalitis; ischaemia;  
KW spongiform encephalopathy; dysexia; age-related memory loss;  
KW Lou Gehring's disease; viral infection; HIV; bacterial infection.

XX Human immunodeficiency virus.

XX WO200062067-A1.

XX 19-OCT-2000.

XX 28-FEB-2000; 2000WO-US005097.

XX 28-FEB-1999; 99US-0122757P.

XX 29-AUG-1999; 99US-0151291P.

XX (UNIW ) UNIV WASHINGTON.

XX Dowdy SF;

XX WPI; 2000-647439/62.

XX Fusion molecules comprising protein transduction domains and therapeutic  
XX agents, useful for treating e.g. Alzheimer's and Parkinson's diseases,  
XX dementia and epilepsy.

XX Disclosure; Page 50; 191pp; English.

XX The invention relates to a novel fusion molecule comprising at least one  
CC protein transduction domain (PTD) and at least one linked molecule, where  
CC the linked molecule has therapeutic or prophylactic activity against a  
CC medical condition. The invention also relates to methods of drug  
CC discovery in which the test compound is linked to a suitable transducing  
CC protein and introduced to a cell; a method of killing resistant  
CC microorganisms using a suitable fusion molecule; a mammal comprising a  
CC covalently linked fusion molecule; and a mammal adapted for experimental  
CC use in which at least one transduction molecule has been transduced into  
CC essentially all the cells of the mammal. The fusion molecule is used to  
CC deliver a therapeutic agent to a mammal, especially a human. The linked  
CC molecule may be a vaccine, an anti-infective drug, a cardiovascular drug,  
CC an antitumour drug, an analgesic, an anti-inflammatory, a diagnostic  
CC marker or a drug for the treatment or prevention of a central or

CC peripheral nervous system disorder. The central nervous system (CNS)  
 CC disorder is especially Alzheimer's disease, Parkinson's disease,  
 CC Huntington's disease, and also includes pre-senile dementia, epilepsy and  
 CC seizures, compulsive behaviour, meningitis (including viral and bacterial  
 CC meningitis), encephalitis, ischaemia, scrapie (or related spongiform  
 CC encephalopathies), dyslexia, age-related memory loss or Lou Gehring's  
 CC disease. Fusion molecules can also be used to kill virally infected  
 CC cells, especially those infected with HIV. The vaccines are used to treat  
 CC or prevent bacterial or viral infections. The methods are a highly  
 CC effective means for transducing a molecule into an entire mammal or into  
 CC specific cells, tissues, organs and systems within it. They also overcome  
 CC bioavailability problems that are associated with many therapeutic agents  
 CC (e.g., large molecular size, hydrophobicity, hydrophilicity, biological  
 CC resistance), by providing efficient transduction of the target cell. The  
 CC present sequence represents a protein transduction domain used in the  
 CC invention

XX  
 SQ Sequence 11 AA;

Query Match 100.0%; Score 58; DB 3; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.013;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YGRKKRQRRR 11  
 |||||  
 Db 1 YGRKKRQRRR 11

RESULT 11  
 AAB03932  
 ID AAB03932 standard; peptide; 11 AA.  
 AC AAB03932;  
 XX  
 DT 12-SEP-2003 (revised)  
 DT 26-FEB-2001 (first entry)  
 XX  
 DE TAT protein transduction domain (internalisation moiety).  
 KW Modulating agent; beta-catenin; hair loss; hair growth; skin;  
 KW exfoliation; Alzheimer's disease; gene transduction;  
 KW cell differentiation; hearing loss; inner ear; hyperacusis; tinnitus;  
 KW hair regeneration.  
 XX  
 OS Human immunodeficiency virus; (HIV).  
 XX  
 PN WO200059939-A1.  
 XX  
 PD 12-OCT-2000.  
 XX  
 PF 04-APR-2000; 2000WO-US009174.  
 XX  
 PR 05-APR-1999; 99US-00288373.  
 XX  
 PA (ADHE-) ADHEREX TECHNOLOGIES INC.  
 XX  
 PI Blaschuk OW, Byers S, Gour BJ;  
 XX  
 DR WPI; 2000-679355/66.  
 XX  
 PT Modulating agents for inhibiting degradation of cytoplasmic beta-catenin,  
 PT used for e.g. stimulating hair growth or reducing hair loss, inhibiting  
 PT development of Alzheimer's disease, comprise internalization moiety and  
 PT amino acid sequence.  
 XX  
 PS Claim 4; Page 46; 49pp; English.  
 XX  
 CC Modulating agents for inhibiting degradation of cytoplasmic beta-catenin  
 CC are described. The modulating agent comprises an internalisation moiety  
 CC and one or more of an amino acid sequence SVDS (PO.4)GHS (PO.4)G, or a  
 CC peptide analogue or peptidomimetic of the amino acid sequence. The  
 CC modulating agents are useful for the manufacture of a medicament for  
 CC stimulating hair growth or reducing hair loss, stimulating skin

CC exfoliation, and inhibiting the development of Alzheimer's disease. They  
 CC may also be used to increase the beta-catenin level in a cell, to  
 CC stimulate activation of gene transcription in a cell, and to stimulate  
 CC cell differentiation. They may further be used to ameliorate hearing loss  
 CC resulting from a variety of inner ear disorders, such as hyperacusis and  
 CC tinnitus, through regeneration of hair cells of the inner ear. (Updated  
 CC on 12-SEP-2003 to standardise OS field)

XX  
 SQ Sequence 11 AA;

Query Match 100.0%; Score 58; DB 3; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.013;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YGRKKRQRRR 11  
 |||||  
 Db 1 YGRKKRQRRR 11

RESULT 12  
 AAB71757  
 ID AAB71757 standard; protein; 11 AA.  
 XX  
 AC AAB71757;  
 XX  
 DT 02-MAY-2001 (first entry)  
 XX  
 DE HIV TAT protein transduction domain.  
 KW HIV; human immunodeficiency virus; protein transduction domain; NTR3;  
 KW tumour necrosis factor receptor; TNF receptor; anti-HIV; antianaemic;  
 KW immunosuppressive; antidiabetic; antiviral; antibacterial; cytostatic;  
 KW neuroprotective; antiinflammatory; anorectic; vasotropic; antirheumatoid;  
 KW antiarthritic; cerebroprotective; tuberculostatic; gene therapy; cancer;  
 KW blood disorder; brain disorder; autoimmune disease; infection.  
 XX  
 OS Human immunodeficiency virus.  
 XX  
 PN WO200110908-A1.  
 XX  
 PD 15-FEB-2001.  
 XX  
 PF 02-AUG-2000; 2000WO-US021287.  
 XX  
 PR 04-AUG-1999; 99US-0147297P.  
 XX  
 PA (AMGE-) AMGEN INC.  
 XX  
 PI Hsu H;  
 XX  
 DR WPI; 2001-191521/19.  
 XX  
 PT New tumor necrosis factor receptor, NTR3, useful for treating cancers,  
 PT stroke, anemia, obesity, rheumatoid arthritis and transplantation  
 PT rejection.  
 XX  
 PS Disclosure; Page 85; 135pp; English.  
 XX  
 CC The present sequence can be used to internalise proteins into a cell by  
 CC targeting the lipid bi-layer component of the cell membrane. The sequence  
 CC may be fused to an NTR3 antagonist and administered intracellularly to  
 CC inhibit the activity of the NTR3 molecule. NTR3 is a tumour necrosis  
 CC factor (TNF) receptor. NTR3 polynucleotides and polypeptides are useful  
 CC for treating diseases such as acquired-immunodeficiency syndrome (AIDS),  
 CC anaemia, autoimmune diseases, cachexia, cancer, cerebral malaria,  
 CC sick syndrome, haemorrhagic shock, hepatitis, insulin resistance,  
 CC leprosy, leukaemia, meningitis, multiple sclerosis, myocardial ischaemia,  
 CC obesity, rejection of transplanted organs, rheumatoid arthritis, septic  
 CC shock syndrome, stroke, adult respiratory distress syndrome (ARDS),  
 CC tuberculosis, and a number of viral diseases. The NTR3 polypeptide is  
 CC useful for identifying or developing new (ant)agonists of NTR3. It may be  
 CC used as an immunogen to which antibodies may be raised. NTR3 nucleic acid

CC molecules may be useful as hybridisation probes in diagnostic assays to  
 CC test, either qualitatively or quantitatively, for the presence of an NTR3  
 CC DNA or corresponding RNA in mammalian tissue or bodily fluid samples  
 XX Sequence 11 AA;

Query Match 100.0%; Score 58; DB 4; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.013;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YGRKKRQRRR 11  
 |||||  
 DB 1 YGRKKRQRRR 11

RESULT 13  
 AAB71756  
 ID AAB71756 standard; protein; 11 AA.

XX AAB71756;

AC AAB71756;

XX 02-MAY-2001 (first entry)

XX NTR3 derived peptide.

XX Human; NTR3; tumour necrosis factor receptor; TNF receptor; anti-HIV;  
 KW antianaemic; immunosuppressive; antidiabetic; antiviral; antibacterial;  
 KW cytosolic; neuroprotective; antiinflammatory; anorectic; vasotropic;  
 KW antirheumatoid; antiarthritic; cerebroprotective; tuberculostatic;  
 KW gene therapy; cancer; blood disorder; brain disorder; autoimmune disease;  
 KW infection.

XX Synthetic.

XX WO200110908-A1.

XX 15-FEB-2001.

XX 02-AUG-2000; 2000WO-US021287.

XX 04-AUG-1999; 99US-0147297P.

XX (AMGE-) AMGEN INC.

XX Hsu H;

XX WPI; 2001-191521/19.

XX New tumor necrosis factor receptor, NTR3, useful for treating cancers,  
 PT stroke, anemia, obesity, rheumatoid arthritis and transplantation  
 PT rejection.

XX Disclosure; Page 131; 135pp; English.

XX The present sequence is a peptide within which conservative and non-  
 CC conservative amino acid substitutions may be made to generate analogues  
 CC of the tumour necrosis factor (TNF) receptor polypeptide NTR3. NTR3  
 CC polynucleotides and polypeptides are useful for treating diseases such as  
 CC acquired-immunodeficiency syndrome (AIDS), anaemia, autoimmune diseases,  
 CC cachexia, cancer, cerebral malaria, diabetes mellitus, disseminated  
 CC intravascular coagulopathy, erythroid sick syndrome, haemorrhagic shock,  
 CC hepatitis, insulin resistance, leprosy, leukaemia, meningitis, multiple  
 CC sclerosis, myocardial ischaemia, obesity, rejection of transplanted  
 CC organs, rheumatoid arthritis, septic shock syndrome, stroke, adult  
 CC respiratory distress syndrome (ARDS), tuberculosis, and a number of viral  
 CC diseases. The NTR3 polypeptide is useful for identifying or developing  
 CC new (ant)agonists of NTR3. It may be used as an immunogen to which  
 CC antibodies may be raised. NTR3 nucleic acid molecules may be useful as  
 CC hybridisation probes in diagnostic assays to test, either qualitatively  
 CC or quantitatively, for the presence of an NTR3 DNA or corresponding RNA  
 CC in mammalian tissue or bodily fluid samples

XX Sequence 11 AA;

Query Match 100.0%; Score 58; DB 4; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.013;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YGRKKRQRRR 11  
 |||||  
 DB 1 YGRKKRQRRR 11

RESULT 14  
 AAB60006  
 ID AAB60006 standard; peptide; 11 AA.

XX AAB60006;

XX 05-NOV-2001 (first entry)

XX Internalising peptide SEQ ID NO: 21.

XX Internalising peptide; transport; apoptosis; arthritis; cancer;  
 KW stem cell; cell differentiation; immune response stimulation;  
 KW HIV vaccine.

XX Synthetic.

XX WO200115511-A2.

XX 08-MAR-2001.

XX 31-AUG-2000; 2000WO-US024034.

XX 01-SEP-1999; 99US-0151980P.

XX 13-MAR-2000; 2000US-0188944P.

XX (UYPI-) UNIV PITTSBURGH.

XX Robbins PD, Mi Z, Frizzell R, Glorioso JC, Gambotto A;

XX WPI; 2001-273309/28.

XX Peptides that facilitate uptake and cytoplasmic and/or nuclear transport  
 PT of proteins, DNA and viruses, useful, e.g. for facilitating uptake of  
 PT antigens in immunogenic compositions.

XX Example 4; Page 117; 129pp; English.

XX The present invention provides the sequences of 75 peptides which  
 CC facilitate the uptake and transport of viruses, proteins and nucleic  
 CC acids. These internalising peptides can be used for transport into the  
 CC cytoplasm or the nucleus. They are useful for facilitating uptake into  
 CC the cell, inducing apoptosis, for example in the treatment of arthritis  
 CC and cancer, to expand a population of stem cells or differentiated cells,  
 CC to stimulate cell differentiation, facilitate the integration of AAV into  
 CC the genome of a cell, and to stimulate an immune response, for example in  
 CC the case of a HIV vaccine. The present sequence is one of the peptides of  
 CC the invention

XX Sequence 11 AA;

Query Match 100.0%; Score 58; DB 4; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.013;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YGRKKRQRRR 11  
 |||||  
 DB 1 YGRKKRQRRR 11

RESULT 15  
 AAE05268  
 ID AAE05268 standard; peptide; 11 AA.

XX

AC AAE05268;  
 XX  
 DT 12-SEP-2001 (first entry)  
 XX  
 DE Human immunodeficiency virus (HIV) TAT peptide.  
 XX  
 KW DNA recombinase domain; protein transduction domain; PTD;  
 KW gene alteration; TAT; fusion protein; Human immunodeficiency virus; HIV.  
 XX  
 OS Human immunodeficiency virus.  
 XX  
 FN WO200149832-A2.  
 XX  
 PD 12-JUL-2001.  
 XX  
 XX 05-JAN-2001; 2001WO-EP000060.  
 PF  
 XX 07-JAN-2000; 2000EP-00100351.  
 PR  
 XX 10-NOV-2000; 2000EP-00124595.  
 XX  
 PA (ARTE-) ARTEMIS PHARM GMBH.  
 XX  
 PI Schwenk F;  
 XX  
 DR WPI; 2001-441873/47.  
 DR N-PSDB; AAD09261.  
 XX  
 PT Using site-specific DNA recombinase domain/protein transduction domain  
 PT fusion proteins for inducing target gene alterations in organisms or cell  
 PT cultures.  
 XX  
 PS Claim 5; Page 70; 85pp; English.  
 XX  
 CC The present invention relates to use of fusion proteins comprising a site  
 CC -specific DNA recombinase domain e.g. Cre and a protein transduction  
 CC domain (PTD) e.g. the Human immunodeficiency virus (HIV) derived TAT  
 CC peptide, for preparing an agent for inducing target gene alterations in a  
 CC living organism or cell culture. The present invention also provides a  
 CC method for inducing gene alterations in living organisms using the fusion  
 CC proteins of the invention. The present sequence is HIV TAT peptide  
 XX  
 SQ Sequence 11 AA;  
 Query Match 100.0%; Score 58; DB 4; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.013;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 YGRKKRQRR 11  
 |||||  
 Db 1 YGRKKRQRR 11

Search completed: March 31, 2005, 00:28:21  
 Job time : 12.7 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2005, 00:19:36 ; Search time 129.942 Seconds  
(without alignments)  
918.211 Million cell updates/sec

Title: US-10-049-822A-2  
Perfect score: 1218  
Sequence: 1 MSQSNRELVDLFLSYKLSQK.....FLTGMTVAGVLLGLSLFSRK 233

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt\_03.\*  
1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1218	100.0	233	1	BCLX RAT
2	1210	99.3	233	1	BCLX MOUSE
3	1190	97.7	233	1	BCLX HUMAN
4	1190	97.7	233	2	Q76LH7
5	1184	97.2	233	2	Q9N1A2
6	1182	97.0	233	2	Q9MZS7
7	1181	97.0	233	1	BCLX_PIG
8	1178	96.7	233	2	Q8SQ42
9	1174	96.4	233	2	Q9MYW4
10	1132	92.9	217	2	Q9N3S5
11	987	81.0	284	2	Q7TS62
12	979	80.4	188	2	Q9CWX2
13	979	80.4	235	2	Q35843
14	963	79.1	188	2	Q9H1R6
15	929	76.3	180	2	Q9BDD5
16	917	75.3	229	1	BCLX_CHICK
17	910	74.7	180	2	Q9BDX7
18	901	74.0	219	2	Q99N36
19	824.5	67.7	170	2	Q9WU15
20	623	51.1	125	2	Q9H1R5
21	615.5	50.5	238	2	Q90Z98
22	529	43.4	201	2	Q6GL15
23	516.5	42.4	204	1	ARL1_XENLA
24	515.5	42.3	204	2	Q90ZH2
25	497	40.8	233	1	BCL2_CHICK
26	474	38.9	239	2	Q75SV7
27	473.5	38.9	229	1	BCL2_BOVIN
28	467.5	38.4	236	1	BCL2_RAT
29	465.5	38.2	236	1	BCL2_CRIOLO
30	465.5	38.2	236	2	Q923R6
31	464	38.1	239	1	BCL2_HUMAN

32	462.5	38.0	236	2	Q6R755
33	462	37.9	235	2	Q81008
34	459.5	37.7	236	1	BCL2_MOUSE
35	458.5	37.6	236	2	Q7TSN8
36	458.5	37.6	236	2	Q8BOK4
37	439.5	36.1	188	2	Q6GP82
38	434.5	35.7	185	2	Q8MJ81
39	434	35.6	228	1	ARL1_XENLA
40	433.5	35.6	193	2	Q88996
41	433.5	35.6	219	2	Q7TS60
42	428.5	35.2	192	2	Q6A093
43	428.5	35.2	193	1	BCLW_HUMAN
44	428.5	35.2	193	1	BCLW_MOUSE
45	421	34.6	199	2	Q6NTH7

## ALIGNMENTS

RESULT 1  
BCLX RAT  
ID BCLX RAT STANDARD; PRT; 233 AA.  
AC P33563: P70613; P70614; Q62678; Q62836; Q64087; Q64128;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE Apoptosis regulator Bcl-X (Bcl-2-like 1 protein).  
GN Name=Bcl2l1; Synonyms=Bclx, B1c2l;  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORMS X(L) AND X(S)).  
RC TISSUE=Brain;  
RA Michaelidis T.M.;  
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RA Wesselingh S.L., David G.L., Choi S., Veliuona M., Hardwick J.M.;  
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A. (ISOFORMS X(L) AND X(BETA)).  
RC TISSUE=Thymus;  
RX MEDLINE=96278736; PubMed=8662675; DOI=10.1074/jbc.271.22.13258;  
RA Shiraiwa N., Inohara N., Okada S., Yuzaki M., Shoji S.-I., Ohta S.;  
RT "An additional form of rat Bcl-x, Bcl-xbeta, generated by an unspliced RNA, promotes apoptosis in promyeloid cells.";  
RL J. Biol. Chem. 271:13258-13265 (1996).  
RN [4]  
RP SEQUENCE FROM N.A. (ISOFORMS X(L) AND X(S)).  
RC STRAIN=Sprague-Dawley; TISSUE=Ovary;  
RX MEDLINE=95129487; PubMed=7828536; DOI=10.1210/en.136.1.232;  
RA Tilly J.L., Tilly K.I., Kenton M.L., Johnson A.L.;  
RT "Expression of members of the bcl-2 gene family in the immature rat ovary: equine chorionic gonadotropin-mediated inhibition of granulosa cell apoptosis is associated with decreased bax and constitutive bcl-2 and bcl-x-long messenger ribonucleic acid levels.";  
RL Endocrinology 136:232-241 (1995).  
RN [5]  
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).  
RX MEDLINE=98010630; PubMed=9346936; DOI=10.1074/jbc.272.44.27886;  
RA Arimoto M., Kunishima N., Inohara N., Ishibashi Y., Ohta S., Morikawa K.;  
RT "Crystal structure of rat Bcl-xL. Implications for the function of the Bcl-2 protein family.";  
RL J. Biol. Chem. 272:27886-27892 (1997).  
CC -1- FUNCTION: Potent inhibitor of cell death. Isoform Bcl-X(L) anti-apoptotic activity is inhibited by association with SIVA isoform 1. Inhibits activation of caspases (By similarity). Appears to regulate cell death by blocking the voltage-dependent anion channel (VDAC) by binding to it and preventing the release of the

caspase activator, cytochrome c, from the mitochondrial membrane.  
 The Bcl-X(S) and Bcl-X(beta) isoforms promote apoptosis.  
 -! SUBUNIT: Bcl-X(L) forms heterodimers with BAX, BAK and Bcl-2 (By similarity). Heterodimerization with BAX does not seem to be required for anti-apoptotic activity (By similarity). Isoform Bcl-X(L) binds to Siva isoform 1 (By similarity).  
 -! SUBCELLULAR LOCATION: Mitochondrial membranes and perinuclear envelope (By similarity).  
 -! ALTERNATIVE PRODUCTS:  
 Event-Alternative splicing; Named isoforms=3;  
 Name=Bcl-X(L);  
 IsoId=P53563-1; Sequence=Displayed;  
 Name=Bcl-X(S);  
 IsoId=P53563-2; Sequence=VSP\_000520;  
 Name=Bcl-X(beta);  
 IsoId=P53563-3; Sequence=VSP\_000521;  
 -! TISSUE SPECIFICITY: Expressed in most tissues. Bcl-X(beta) is specifically expressed in cerebellum, heart, and thymus. In the ovary, the predominant form is Bcl-X(L), with a small but detectable level of Bcl-X(S).  
 -! DOMAIN: The BH4 domain is required for anti-apoptotic activity. The BH1 and BH2 domains are required for both heterodimerization with other Bcl-2 family members and for repression of cell death.  
 -! PTM: Proteolytically cleaved by caspases during apoptosis. The cleaved protein, lacking the BH4 domain, has pro-apoptotic activity (By similarity).  
 -! SIMILARITY: Belongs to the Bcl-2 family.  
 -! SIMILARITY: Contains 1 Bcl-2 homology 1 (BH1) domain.  
 -! SIMILARITY: Contains 1 Bcl-2 homology 2 (BH2) domain.  
 -! SIMILARITY: Contains 1 Bcl-2 homology 3 (BH3) domain.  
 -! SIMILARITY: Contains 1 Bcl-2 homology 4 (BH4) domain.  
 -----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 EMBL; X82537; CAA57886.1; -;  
 DR EMBL; X82537; CAA57887.1; -;  
 DR EMBL; U10579; AAA19257.1; -;  
 DR EMBL; U72350; AAB17353.1; -;  
 DR EMBL; U72349; AAB17352.1; -;  
 DR EMBL; U34963; AAA77686.1; -;  
 DR EMBL; S76513; AAC60701.2; ALT\_INIT.  
 DR EMBL; S78284; AAC60702.1; -;  
 DR PIR; I67431; I67431.  
 DR PIR; S51761; S51761.  
 DR PDB; 1AF3; X-ray; @=1-196.  
 DR RGD; 2200; Bcl2l1.  
 DR InterPro; IPR000712; Bcl2\_BH.  
 DR InterPro; IPR003093; Bcl2\_BH4.  
 DR InterPro; IPR002475; Bcl2\_family.  
 DR InterPro; IPR004725; Bcl2\_reg.  
 DR Pfam; PF00452; Bcl-2; 1.  
 DR Pfam; PF02180; BH4; 1.  
 DR TIGRFAMs; TIGR00865; bcl-2; 1.  
 DR PROSITE; PS50062; BCL2\_FAMILY; 1.  
 DR PROSITE; PS01080; BH1; 1.  
 DR PROSITE; PS01258; BH2; 1.  
 DR PROSITE; PS01259; BH3; 1.  
 DR PROSITE; PS01260; BH4; 1; 1.  
 DR PROSITE; PS50063; BH4 2; 1.  
 KW 3D-structure; Alternative splicing; Apoptosis; Mitochondrion;  
 KW Transmembrane.  
 FT DOMAIN 4 24 BH4.  
 FT DOMAIN 86 100 BH3.  
 FT DOMAIN 129 148 BH1.  
 FT DOMAIN 180 195 BH2.  
 FT TRANSMEM 210 226 Potential.  
 FT VARSPPLIC 126 188 Missing (in isoform Bcl-X(S)).

/FTid=VSP\_000520.  
 DTFVDLYGNNAAESRKGOERFNRWFLTGMTVAGVLLGSL  
 PSRK -> VRTPLVCPPLVCLSSVETPCFPHSGMVVED  
 IDSGDIPGLL (in isoform Bcl-X(beta)).  
 /FTid=VSP\_000521.  
 R -> Q (in Ref. 1).  
 F -> S (in Ref. 2).  
 A -> E (in Ref. 2).  
 I -> L (in Ref. 4).  
 A -> V (in Ref. 4).  
 FF -> SS (in Ref. 4).  
 A -> T (in Ref. 4).  
 A -> P (in Ref. 4).  
 FT VARSPPLIC 189 233  
 FT CONFLICT 6 6  
 FT CONFLICT 12 12  
 FT CONFLICT 64 64  
 FT CONFLICT 81 81  
 FT CONFLICT 119 119  
 FT CONFLICT 143 144  
 FT CONFLICT 199 199  
 FT CONFLICT 201 201  
 FT HELIX 4 19  
 FT TURN 20 21  
 FT TURN 25 28  
 FT TURN 82 83  
 FT HELIX 84 100  
 FT HELIX 106 112  
 FT TURN 116 117  
 FT HELIX 120 127  
 FT HELIX 128 131  
 FT TURN 132 133  
 FT HELIX 137 156  
 FT TURN 157 158  
 FT TURN 160 161  
 FT HELIX 162 177  
 FT TURN 178 178  
 FT HELIX 179 186  
 FT TURN 185 186  
 FT HELIX 187 195  
 SQ SEQUENCE 233 AA; 26158 MW; 2B62B6C3864BC8F CRC64;  
 Query Match 100.0%; Score 1218; DB 1; Length 233;  
 Best Local Similarity 100.0%; Pred. No. 4.4e-97;  
 Matches 233; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 MSQNRRLVDFLSYKLSQKGYSSQFSDEENTEAPEETEPEPEPSAINGNPSWHLA 60  
 Db 1 MSQNRRLVDFLSYKLSQKGYSSQFSDEENTEAPEETEPEPEPSAINGNPSWHLA 60  
 Qy 61 DSPAVNGATGHSSSLDAREVTPMAAVKQALREAGDEPELRYRRAFSDLTSQLHITPGTAY 120  
 Db 61 DSPAVNGATGHSSSLDAREVTPMAAVKQALREAGDEPELRYRRAFSDLTSQLHITPGTAY 120  
 Qy 121 QSPFQVNNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLVSRIASWMTYLNHLEP 180  
 Db 121 QSPFQVNNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLVSRIASWMTYLNHLEP 180  
 Qy 181 WIQENGGWDTFVDLYGNNAAESRKGOERFNRWFLTGMTVAGVLLGSLPSRK 233  
 Db 181 WIQENGGWDTFVDLYGNNAAESRKGOERFNRWFLTGMTVAGVLLGSLPSRK 233  
 RESULT 2  
 BCLX\_MOUSE  
 ID\_BCLX\_MOUSE STANDARD; PRT; 233 AA.  
 AC Q64373; Q35844; Q60657; Q60658; Q61338;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Apoptosis regulator Bcl-X (Bcl-2-like 1 protein).  
 GN Name=Bcl2l1; Synonyms=Bcl12l; Bclx;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_taxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM X(L)).  
 RC STRAIN=2A4B;  
 RA Kamesaki H., Michaud G.Y., Takatsu K., Okuma M.;  
 RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.  
 RN [2]

RP SEQUENCE FROM N.A. (ISOFORMS X(L) AND X(BETA)).

RC STRAIN=C57BL/6; TISSUE=Brain;

RX MEDLINE=9531139; PubMed=7607090;

RA Gonzalez-Carbia M., Perez-Ballester R., Ding L., Duan L., Boise L.H.,

RA Thompson C.B., Nunez G.;

FT "bcl-Xl is the major bcl-x mRNA form expressed during murine

RT development and its product localizes to mitochondria.";

RL Development 120:3033-3042(1994).

RN (3)

RP SEQUENCE FROM N.A. (ISOFORMS X(L); X(S) AND X(Delta-TM)).

RC TISSUE=Pre-B cell;

RX MEDLINE=95052604; PubMed=7963517;

RA Fang W., Rivard J.J., Mueller D.L., Behrens T.W.;

RT "Cloning and molecular characterization of mouse bcl-x in B and T

RL lymphocytes.";

RN J. Immunol. 153:4388-4398(1994).

RN (4)

RP SEQUENCE FROM N.A. (ISOFORMS X(L) AND X(BETA)).

RC STRAIN=C57BL/6 X CBA; TISSUE=Thymus;

RX MEDLINE=98051053; PubMed=9390687; DOI=10.1016/S1074-7613(00)80384-2;

RA Yang X.-F., Weber G.F., Cantor H.;

RT "A novel Bcl-x isoform connected to the T cell receptor regulates

RL apoptosis in T cells.";

RN Immunity 7:629-639(1997).

RN (5)

RP SEQUENCE FROM N.A.

RX MEDLINE=97289584; PubMed=9144489;

RA Grillo D.A., Gonzalez-Garcia M., Ekhterae D., Duan L., Inohara N.,

RA Ohta S., Seldin M.F., Nunez G.;

RT "Genomic organization, promoter region analysis, and chromosome

RL localization of the mouse bcl-x gene.";

RN J. Immunol. 158:4750-4757(1997).

CC -!- FUNCTION: Potent inhibitor of cell death. Isoform Bcl-X(L) anti-

CC apoptotic activity is inhibited by association with SIVA isoform

CC 1. Inhibits activation of caspases (By similarity). Appears to

CC regulate cell death by binding and blocking the voltage-dependent

CC anion channel (VDAC), thus preventing the release of the caspase

CC activator, cytochrome c, from the mitochondrial membrane. The Bcl-

CC X(S) isoform promotes apoptosis.

CC -!- SUBUNIT: Bcl-X(L) forms heterodimers with BAX, BAK and Bcl-2 (By

CC similarity). Heterodimerization with BAX does not seem to be

CC required for anti-apoptotic activity (By similarity). Isoform Bcl-

CC X(L) binds to Siva isoform 1 (By similarity).

CC -!- SUBCELLULAR LOCATION: Mitochondrial membranes and perinuclear

CC envelope for Bcl-X(L). Cytoplasmic for Bcl-X(Delta-TM).

CC -!- ALTERNATIVE PRODUCTS:

CC Event-Alternative splicing; Named isoforms=4;

CC Name=BCL-X(L);

CC IsoId=Q64373-1; Sequence=Displayed;

CC Name=BCL-X(S);

CC IsoId=Q64373-2; Sequence=VSP\_000517;

CC Name=BCL-X(beta);

CC IsoId=Q64373-3; Sequence=VSP\_000518;

CC Name=BCL-X(Delta-TM);

CC IsoId=Q64373-4; Sequence=VSP\_000519;

CC -!- TISSUE SPECIFICITY: Widely expressed, with highest levels in the

CC brain, thymus, bone marrow, and kidney. Bcl-X(L) and Bcl-X(Delta-

CC TM) expression is enhanced in B and T lymphocytes that have been

CC activated.

CC -!- DEVELOPMENTAL STAGE: Bcl-X(beta) is expressed in both embryonal

CC and postnatal tissues, whereas Bcl-X(L) is predominantly found in

CC postnatal tissues.

CC -!- DOMAIN: The BH4 domain is required for anti-apoptotic activity.

CC The BH1 and BH2 domains are required for both heterodimerization

CC with other Bcl-2 family members and for repression of cell death.

CC -!- PTM: Proteolytically cleaved by caspases during apoptosis (By

CC similarity). The cleaved protein, lacking the BH4 domain, has pro-

CC apoptotic activity (By similarity).

CC -!- SIMILARITY: Belongs to the Bcl-2 family.

CC -!- SIMILARITY: Contains 1 Bcl-2 homology 1 (BH1) domain.

CC -!- SIMILARITY: Contains 1 Bcl-2 homology 2 (BH2) domain.

CC -!- SIMILARITY: Contains 1 Bcl-2 homology 3 (BH3) domain.

CC -!- SIMILARITY: Contains 1 Bcl-2 homology 4 (BH4) domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; X83574; CAA58557.1; -

DR EMBL; L35049; AAA51039.1; -

DR EMBL; L35048; AAA51040.1; -

DR EMBL; U10102; AAA82174.1; -

DR EMBL; U10101; AAA82173.1; -

DR EMBL; U10100; AAA82172.1; -

DR EMBL; U51278; AAC53459.1; -

DR EMBL; U51279; AAC53460.1; -

DR EMBL; U78031; AAB96881.1; -

DR EMBL; U78030; AAB96881.1; JOINED.

DR PIR; I49055; I49055.

DR PIR; I49057; I49057.

DR HSSP; P53563; 1AF3.

DR MGD; MGI:88139; Bcl2l.

DR GO; GO:0005515; F:protein binding; IPI.

DR GO; GO:0006916; P:anti-apoptosis; IDA.

DR GO; GO:0006915; P:apoptosis; IDA.

DR InterPro; IPR000712; Bcl2\_BH.

DR InterPro; IPR003093; Bcl2\_BH4.

DR InterPro; IPR002475; BCL2\_family.

DR InterPro; IPR004725; Bcl2\_reg.

DR Pfam; PF00452; Bcl-2; 1.

DR Pfam; PF02180; BH4; 1.

DR TIGRPFAMs; TIGR00865; bcl-2; 1.

DR PROSITE; PS50062; BCL2\_FAMILY; 1.

DR PROSITE; PS01080; BH1; 1.

DR PROSITE; PS01258; BH2; 1.

DR PROSITE; PS01259; BH3; 1.

DR PROSITE; PS01260; BH4; 1.

DR PROSITE; PS50063; BH4; 2; 1.

DR PROSITE; PS50063; BH4; 2; 1.

KW Alternative splicing; Apoptosis; Mitochondrion; Transmembrane.

FT DOMAIN 4 24 BH4.

FT DOMAIN 86 -100 BH3.

FT DOMAIN 129 148 BH1.

FT DOMAIN 180 195 BH2.

FT TRANSMEM 210 226 Potential.

FT VARSPPLIC 126 188 Missing (in isoform BCL-X(S)).

FT VARSPPLIC 189 233 /FTId=VSP\_000517.

FT VARSPPLIC 194 233 DTFVLDYGNNAASRRKQERFNRWFLTGMTVAGVVLGSL

FT VARSPPLIC 194 233 FSRK -> VRTPLVCPPLACVSLCEHP (in isoform

FT VARSPPLIC 194 233 LYGNNAASRRKQERFNRWFLTGMTVAGVVLGSLFSRK

FT VARSPPLIC 194 233 -> GHDCWCGSAGLTQLQSEVTRH (in isoform BCL-

FT VARSPPLIC 194 233 X(Delta-TM)).

FT VARSPPLIC 194 233 /FTId=VSP\_000518.

FT VARSPPLIC 194 233 QER -> KEG (in Ref. 4; AAC53459).

FT VARSPPLIC 194 233 CONFLICT 207 209

FT VARSPPLIC 194 233 SEQUENCE 233 AA; 26132 MW; 24D2AC79887E072E CRC64;

Query Match 99.3%; Score 1210; DB 1; Length 233;

Best Local Similarity 99.6%; Pred. No. 2.2e-96;

Matches 232; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MSQSNRELVDVFLSYKLSQKYSQSDVEENTEAPEETEPEETPSAINGPSWHLA 60

Db 1 MSQSNRELVDVFLSYKLSQKYSQSDVEENTEAPEETEPEETPSAINGPSWHLA 60

Qy 61 DSPAVNGATGHSSSLDAREVPMMAVKQALREAGDEFEELRYRRAFSDLTSLHITPGTAY 120

Db 61 DSPAVNGATGHSSSLDAREVPMMAVKQALREAGDEFEELRYRRAFSDLTSLHITPGTAY 120

Qy 121 QSPFQVYNELFRDGVNWRIVAFPSFGALCVESVDKEMQVLSVRIASWMTATYNDHLEP 180

```

Db 121 QSFQVNNELFRDGVNMGRIVAFFSGALCVESVDKQVLVSKIASWMTALNDHLEP 180
Qy 181 WIOENGCHDFTVDLYGNNAASRKQGRFNRWFLTGMTAGVLLGSLFSRK 233
    |||||
Db 181 WIOENGCHDFTVDLYGNNAASRKQGRFNRWFLTGMTAGVLLGSLFSRK 233
    |||||

RESULT 3
BCLX_HUMAN
ID BCLX_HUMAN STANDARD; PRT; 233 AA.
AC Q07817; Q92976;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DE 25-OCT-2004 (Rel. 45, Last annotation update)
DE Apoptosis regulator Bcl-X (Bcl-2-like 1 protein).
GN Name=BCL2L1; Synonyms=BCL2L, BCLX;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS X(L) AND X(S)).
RX MEDLINE=93364977; PubMed=8358789; DOI=10.1016/0092-8674(93)90508-N;
RA Boise L.H., Gonzalez-Garcia M., Postema C.E., Ding L., Lindsten T.,
RA Turka L.A., Mao X., Nunez G., Thompson C.B.;
RT "bcl-x, a bcl-2-related gene that functions as a dominant regulator of
RT apoptotic cell death."
RL Cell 74:597-608 (1993).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM X(BETA)).
RA Inohara N., Ohta S.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM X(L)).
RC TISSUE=Lung;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins L.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Rouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Greenwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [4]
RP MUTAGENESIS OF G1Y-138, AND HETERO-DIMERIZATION.
RX MEDLINE=95372373; PubMed=7644501;
RA Sedlak T.W., Oltvai Z.N., Yang E., Wang K., Boise L.H., Thompson C.B.,
RA Korsmeyer S.J.;
RT "Multiple Bcl-2 family members demonstrate selective dimerizations
RT with Bax."
RL Proc. Natl. Acad. Sci. U.S.A. 92:7834-7838 (1995).
RN [5]
RP MUTAGENESIS OF BH1 AND BH2 DOMAINS.
RX MEDLINE=96170038; PubMed=8596636; DOI=10.1038/379554a0;
RA Cheng E.H.-Y., Levine B., Boise L.H., Thompson C.B., Hardwick J.M.,
RA Korsmeyer S.J.;
RT "Bax-independent inhibition of apoptosis by Bcl-XL."
RL Nature 379:554-556 (1996).
RN [6]
RP INTERACTION WITH SIVA.

```

---

```

RX MEDLINE=22008092; PubMed=12011449; DOI=10.1073/pnas.102182299;
RA Xue L., Chu P., Cheng Y., Sun X., Borthakur A., Ramarao M., Pandey P.,
RA Wu M., Schlossman S.F., Prasad K.V.S.;
RT "Siva-1 binds to and inhibits BCL-X(L)-mediated protection against UV
RT radiation-induced apoptosis."
RL Proc. Natl. Acad. Sci. U.S.A. 99:6925-6930 (2002).
RN [7]
RP STRUCTURE BY NMR OF 1-209.
RX MEDLINE=97172562; PubMed=9020082; DOI=10.1126/science.275.5302.983;
RA Sattler M., Liang H., Nettelsheim D., Meadows R.P., Harlan J.E.,
RA Eberstadt M., Yoon H.S., Shuker S.B., Chang B.S., Minn A.J.,
RA Thompson C.B., Fesik S.W.;
RT "Structure of Bcl-XL-Bak peptide complex: recognition between
RT regulators of apoptosis."
RL Science 275:983-986 (1997).
RN [8]
RX X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS), AND STRUCTURE BY NMR OF 1-209.
RP MEDLINE=96256675; PubMed=8692274; DOI=10.1038/381335a0;
RA Muchmore S.W., Sattler M., Liang H., Meadows R.P., Harlan J.E.,
RA Yoon H.S., Nettelsheim D., Chang B.S., Thompson C.B., Wong S.L.,
RA Ng S.L., Fesik S.W.;
RT "X-ray and NMR structure of human Bcl-XL, an inhibitor of programmed
RT cell death."
RL Nature 381:335-341 (1996).
RN [9]
RP CLEAVAGE BY CASPASES, AND MUTAGENESIS OF ASP-61.
RX MEDLINE=98118550; PubMed=9435230; DOI=10.1073/pnas.95.2.554;
RA Clem R.J., Cheng E.H.-Y., Karp C.L., Kirsch D.G., Ueno K.,
RA Takahashi A., Kastan M.B., Griffin D.E., Earnshaw W.C., Veliuona M.A.,
RA Hardwick J.M.;
RT "Modulation of cell death by Bcl-XL through caspase interaction."
RL Proc. Natl. Acad. Sci. U.S.A. 95:554-559 (1998).
CC -!- FUNCTION: Potent inhibitor of cell death. Isoform Bcl-X(L) anti-
CC apoptotic activity is inhibited by association with SIVA isoform
CC 1. Inhibits activation of caspases (By similarity). Appears to
CC regulate cell death by blocking the voltage-dependent anion
CC channel (VDAC) by binding to it and preventing the release of the
CC caspase activator, cytochrome c, from the mitochondrial membrane.
CC The Bcl-X(S) isoform promotes apoptosis.
CC -!- SUBUNIT: Bcl-X(L) forms heterodimers with BAX, BAK and Bcl-2.
CC Heterodimerization with BAX does not seem to be required for anti-
CC apoptotic activity. Isoform Bcl-X(L) binds to Siva isoform 1.
CC -!- SUBCELLULAR LOCATION: Mitochondrial membranes and perinuclear
CC envelope (By similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Name=Bcl-X(L);
CC IsoId=Q07817-1; Sequence=VSP_000515;
CC Name=Bcl-X(S);
CC IsoId=Q07817-2; Sequence=VSP_000516;
CC Name=Bcl-X(beta);
CC IsoId=Q07817-3; Sequence=VSP_000516;
CC -!- TISSUE SPECIFICITY: Bcl-X(S) is expressed at high levels in cells
CC that undergo a high rate of turnover, such as developing
CC lymphocytes. In contrast, Bcl-X(L) is found in tissues containing
CC long-lived postmitotic cells, such as adult brain.
CC -!- DOMAIN: The BH4 domain is required for anti-apoptotic activity.
CC The BH1 and BH2 domains are required for both heterodimerization
CC with other Bcl-2 family members and for repression of cell death.
CC -!- PTM: Proteolytically cleaved by caspases during apoptosis. The
CC cleaved protein, lacking the BH4 domain, has pro-apoptotic
CC activity.
CC -!- SIMILARITY: Belongs to the Bcl-2 family.
CC -!- SIMILARITY: Contains 1 Bcl-2 homology 1 (BH1) domain.
CC -!- SIMILARITY: Contains 1 Bcl-2 homology 2 (BH2) domain.
CC -!- SIMILARITY: Contains 1 Bcl-2 homology 3 (BH3) domain.
CC -!- SIMILARITY: Contains 1 Bcl-2 homology 4 (BH4) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC

```

entities requires a license agreement (see <http://www.iesb-sib.ch/announcement/>) or send an email to [license@iesb-sib.ch](mailto:license@iesb-sib.ch).

EMBL; Z23116; CRA80662.1; -  
EMBL; Z23115; CRA80661.1; -  
EMBL; U72398; AAB17354.1; -  
EMBL; BC019307; AAI19307.1; -  
PIR; B47537; B47537.  
PIR; JE0203; JE0203.  
PDB; 1BXL; NMR; A=1-217.  
PDB; 1G5J; NMR; A=1-211.  
PDB; 1G5M; NMR; A=-.  
PDB; 1GJH; NMR; A=-.  
PDB; 1LXL; NMR; @=1-217.  
PDB; 1MAZ; X-ray; @=1-217.  
PDB; 1R2D; X-ray; A=1-218.  
PDB; 1R2E; X-ray; A=1-218.  
PDB; 1R2G; X-ray; A=1-218.  
PDB; 1R2H; X-ray; A=1-218.  
PDB; 1R2I; X-ray; A=1-218.  
IntAct; Q07817; -  
Genew; HGNC:992; BCL2L1.  
H-InvDB; HIX0015713; -  
MIM; 600039; -; C:mitochondrion; TAS.  
GO; GO:0005739; C:mitochondrion; TAS.  
GO; GO:0006916; P:anti-apoptosis; TAS.  
GO; GO:0006915; P:apoptosis; TAS.  
GO; GO:0008637; P:apoptotic mitochondrial changes; TAS.  
GO; GO:0008634; P:negative regulation of survival gene produc. . .; TAS.  
InterPro; IPR000712; Bcl2\_BH.  
InterPro; IPR003093; Bcl2\_BH4.  
InterPro; IPR002475; Bcl2\_family.  
InterPro; IPR004725; Bcl2\_reg.  
Pfam; PF00452; Bcl-2; 1.  
Pfam; PF02180; BH4; 1.  
TIGRFAMs; TIGR00865; bcl-2; 1.  
PROSITE; PS50062; BCL2\_FAMILY; 1.  
PROSITE; PS01080; BH1; 1.  
PROSITE; PS01258; BH2; 1.  
PROSITE; PS01259; BH3; 1.  
PROSITE; PS01260; BH4; 1.  
PROSITE; PS00063; BH4\_2; 1.  
3D-structure; Alternative splicing; Apoptosis; Mitochondrion; Transmembrane.  
DOMAIN 4 24 BH4.  
DOMAIN 86 100 BH3.  
DOMAIN 129 148 BH1.  
DOMAIN 180 195 BH2.  
TRANSMEM 210 226 Potential.  
SITE 61 62 Cleavage (by caspase-1).  
VARSPPLIC 126 188 Missing (in isoform Bcl-X(S)).  
VARSPPLIC 189 233 /FTId=VSP\_000515.  
FT DTFVELYGNNAESRKQGRFNRWFLTGMTVAGVLLGSL  
FT FSRK -> VRTKPLVCPFSLASGQSPRTALLLYFLLCWV  
FT VGDVDS (in isoform Bcl-X(beta)).  
FT /FTId=VSP\_000516.  
FT D->A: No Cleavage by caspase-1 nor by  
FT caspase-3.  
FT FRD->VRA: No heterodimerization with BAX.  
FT VNW->AIL: Loss of anti-apoptotic  
FT activity.  
FT GRI->ELN: Loss of anti-apoptotic  
FT activity.

Query Match 97.7%; Score 1190; DB 1; Length 233;  
Best Local Similarity 97.9%; Pred. No. 1.2e-94;  
Matches 228; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
Qy 1 MSQSNRELVDVFLSYKLSQKGYSMQSFSDVEENRTEAPEETEPEETPSAINGNPSWHLA 60  
Db 1 MSQSNRELVDVFLSYKLSQKGYSMQSFSDVEENRTEAPEETEPEETPSAINGNPSWHLA 60

Qy 61 DSPAVNGATGHSSSLDAREVIPMAAVKQALREAGDEFELRYRRAFSDLTSLQHTTPGTAY 120  
Db 61 DSPAVNGATGHSSSLDAREVIPMAAVKQALREAGDEFELRYRRAFSDLTSLQHTTPGTAY 120  
Qy 121 QSFQVNVNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLSRIASWMTATYLNHLEP 180  
Db 121 QSFQVNVNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLSRIASWMTATYLNHLEP 180  
Qy 181 WIQENGWDITFDVLYGNNAESRKQGRFNRWFLTGMTVAGVLLGSLFSRK 233  
Db 181 WIQENGWDITFDVLYGNNAESRKQGRFNRWFLTGMTVAGVLLGSLFSRK 233

RESULT 4  
Q76LT7 PRELIMINARY; PRT; 233 AA.  
AC Q76LT7;  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE Bcl-XL.  
GN Name=bcl-xL;  
OS Canis familiaris (Dog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
OC NCBI\_TaxID=9615;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Sano J.;  
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB073983; BAB71819.1; -.  
DR HSSP; PS3563; IAP3.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0042981; P:regulation of apoptosis; IEA.  
DR InterPro; IPR00712; Bcl2\_BH.  
DR InterPro; IPR003093; Bcl2\_BH4.  
DR InterPro; IPR002475; Bcl2\_family.  
DR InterPro; IPR004725; Bcl2\_reg.  
DR Pfam; PF00452; Bcl-2; 1.  
DR Pfam; PF02180; BH4; 1.  
DR SMART; SM00337; BCL; 1.  
DR SMART; SM00265; BH4; 1.  
DR TIGRFAMs; TIGR00865; bcl-2; 1.  
DR PROSITE; PS50062; BCL2\_FAMILY; 1.  
DR PROSITE; PS01080; BH1; 1.  
DR PROSITE; PS01258; BH2; 1.  
DR PROSITE; PS01259; BH3; 1.  
DR PROSITE; PS01260; BH4; 1.  
DR PROSITE; PS00063; BH4\_2; 1.  
SQ SEQUENCE 233 AA; 26049 MW; E09D3CDD851AE9BE CRC64;

Query Match 97.7%; Score 1190; DB 2; Length 233;  
Best Local Similarity 97.9%; Pred. No. 1.2e-94;  
Matches 228; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
Qy 1 MSQSNRELVDVFLSYKLSQKGYSMQSFSDVEENRTEAPEETEPEETPSAINGNPSWHLA 60  
Db 1 MSQSNRELVDVFLSYKLSQKGYSMQSFSDVEENRTEAPEETEPEETPSAINGNPSWHLA 60

Qy 61 DSPAVNGATGHSSSLDAREVIPMAAVKQALREAGDEFELRYRRAFSDLTSLQHTTPGTAY 120  
Db 61 DSPAVNGATGHSSSLDAREVIPMAAVKQALREAGDEFELRYRRAFSDLTSLQHTTPGTAY 120  
Qy 121 QSFQVNVNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLSRIASWMTATYLNHLEP 180  
Db 121 QSFQVNVNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLSRIASWMTATYLNHLEP 180  
Qy 181 WIQENGWDITFDVLYGNNAESRKQGRFNRWFLTGMTVAGVLLGSLFSRK 233  
Db 181 WIQENGWDITFDVLYGNNAESRKQGRFNRWFLTGMTVAGVLLGSLFSRK 233

RESULT 5  
Q9N1A2

Query Match 97.7%; Score 1190; DB 1; Length 233;  
Best Local Similarity 97.9%; Pred. No. 1.2e-94;  
Matches 228; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
Qy 1 MSQSNRELVDVFLSYKLSQKGYSMQSFSDVEENRTEAPEETEPEETPSAINGNPSWHLA 60  
Db 1 MSQSNRELVDVFLSYKLSQKGYSMQSFSDVEENRTEAPEETEPEETPSAINGNPSWHLA 60  
Qy 61 DSPAVNGATGHSSSLDAREVIPMAAVKQALREAGDEFELRYRRAFSDLTSLQHTTPGTAY 120

```

ID Q9N1A2 PRELIMINARY; PRT; 233 AA.
AC Q9N1A2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Anti-apoptotic regulator Bcl-xL.
GN Name=bcl-xL;
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_taxid=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
CA Lee T.L., Canty J.M.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF216205; AAF33212.1; -.
DR HSSP; Q07817; IR2D.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR000712; Bcl2_BH.
DR InterPro; IPR003093; Bcl2_BH4.
DR InterPro; IPR002475; Bcl2_family.
DR InterPro; IPR004725; Bcl2_reg.
DR Pfam; PF00452; Bcl-2; 1.
DR Pfam; PF02180; BH4; 1.
DR SMART; SM00337; BCL; 1.
DR SMART; SM00265; BH4; 1.
DR TIGRFAMs; TIGR00865; bcl-2; 1.
DR PROSITE; PS00062; BCL2_FAMILY; 1.
DR PROSITE; PS01080; BH1; 1.
DR PROSITE; PS01258; BH2; 1.
DR PROSITE; PS01259; BH3; 1.
DR PROSITE; PS01260; BH4; 1.
DR PROSITE; PS01260; BH4_1; 1.
DR PROSITE; PS01260; BH4_2; 1.
DR PROSITE; PS00063; BH4_2; 1.
SQ SEQUENCE 233 AA; 26047 MW; 2FA312818B25E17D CRC64;

Query Match 97.2%; Score 1184; DB 2; Length 233;
Best Local Similarity 97.0%; Pred. No. 3.9e-94;
Matches 226; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MSQSNRELVDVFLSYKLSQKGYWSQFSDVEENRTEAPEETEPEPTPSAINGNPSWHLA 60
Db 1 MSQSNRELVDVFLSYKLSQKGYWSQFSDVEENRTEAPEETEPEPTPSAINGNPSWHLA 60

Qy 61 DSPAVNGATGHSSSLDAREVIPMAAVKQALREAGDEFLRYRRAPSDLTSLQHTTPGTAY 120
Db 61 DSPAVNGATGHSSSLDAREVIPMAAVKQALREAGDEFLRYRRAPSDLTSLQHTTPGTAY 120

Qy 121 QSFQVNVNLFPRDGVNMGRIVAFPSFGGALCVESVDKEMQVLVSRISAMATYLNHLEP 180
Db 121 QSFQVNVNLFPRDGVNMGRIVAFPSFGGALCVESVDKEMQVLVSRISAMATYLNHLEP 180

Qy 181 WIQENGWDFTFDLYGNNAAESRKQGRFNRWFLTGMTVAGVLLGSLFSRK 233
Db 181 WIQENGWDFTFDLYGNNAAESRKQGRFNRWFLTGMTVAGVLLGSLFSRK 233

RESULT 6
Q9MZS7 PRELIMINARY; PRT; 233 AA.
AC Q9MZS7;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE Bcl-x long protein.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Caprinae; Ovis.
OX NCBI_taxid=9940;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC TISSUE=Ovary;
RA Murray J.F., Dong Y.B., Leigh A.J., Scaramuzzi R.J., Carter N.D.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF164517; AAF89532.1; -.
DR HSSP; PS3563; IAP3.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0042981; P:regulation of apoptosis; IEA.
DR InterPro; IPR000712; Bcl2_BH.
DR InterPro; IPR003093; Bcl2_BH4.
DR InterPro; IPR002475; Bcl2_family.
DR InterPro; IPR004725; Bcl2_reg.
DR Pfam; PF00452; Bcl-2; 1.
DR Pfam; PF02180; BH4; 1.
DR SMART; SM00337; BCL; 1.
DR SMART; SM00265; BH4; 1.
DR TIGRFAMs; TIGR00865; bcl-2; 1.
DR PROSITE; PS00062; BCL2_FAMILY; 1.
DR PROSITE; PS01080; BH1; 1.
DR PROSITE; PS01258; BH2; 1.
DR PROSITE; PS01259; BH3; 1.
DR PROSITE; PS01260; BH4; 1.
DR PROSITE; PS00063; BH4_2; 1.
SQ SEQUENCE 233 AA; 26134 MW; 012BFA1382762915 CRC64;

Query Match 97.0%; Score 1182; DB 2; Length 233;
Best Local Similarity 97.0%; Pred. No. 5.8e-94;
Matches 226; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MSQSNRELVDVFLSYKLSQKGYWSQFSDVEENRTEAPEETEPEPTPSAINGNPSWHLA 60
Db 1 MSQSNRELVDVFLSYKLSQKGYWSQFSDVEENRTEAPEETEPEPTPSAINGNPSWHLA 60

Qy 61 DSPAVNGATGHSSSLDAREVIPMAAVKQALREAGDEFLRYRRAPSDLTSLQHTTPGTAY 120
Db 61 DSPAVNGATGHSSSLDAREVIPMAAVKQALREAGDEFLRYRRAPSDLTSLQHTTPGTAY 120

Qy 121 QSFQVNVNLFPRDGVNMGRIVAFPSFGGALCVESVDKEMQVLVSRISAMATYLNHLEP 180
Db 121 QSFQVNVNLFPRDGVNMGRIVAFPSFGGALCVESVDKEMQVLVSRISAMATYLNHLEP 180

Qy 181 WIQENGWDFTFDLYGNNAAESRKQGRFNRWFLTGMTVAGVLLGSLFSRK 233
Db 181 WIQENGWDFTFDLYGNNAAESRKQGRFNRWFLTGMTVAGVLLGSLFSRK 233

RESULT 7
BCLX_PIG STANDARD; PRT; 233 AA.
AC 077737;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Apoptosis regulator Bcl-X (Bcl-2-like 1 protein).
GN Name=BCL2L1; Synonyms=BCLX, BLC2L;
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_taxid=9823;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99171363; PubMed=10072723; DOI=10.1006/jmcc.1998.0855;
RA Bartling B., Hoffmann J., Holtz J., Schulz R., Heusch G., Damer D.;
RT "Quantification of cardioprotective gene expression in porcine short-term hibernating myocardium.";
RL J. Mol. Cell. Cardiol. 31:147-158(1999).
CC -I- FUNCTION: Potent inhibitor of cell death. Isoform Bcl-X(L) anti-apoptotic activity is inhibited by association with SIVA isoform 1. Inhibits activation of caspases (By similarity). Appears to regulate cell death by blocking the voltage-dependent anion channel (VDAC) by binding to it and preventing the release of the caspase activator, cytochrome c, from the mitochondrial membrane.
CC -I- SUBUNIT: Bcl-X(L) forms heterodimers with BAX, BAK and Bcl-2 (By similarity). Heterodimerization with BAX does not seem to be

```

required for anti-apoptotic activity (By similarity). Isoform Bcl-Xl1 binds to Siva isoform 1 (By similarity).

-!- SUBCELLULAR LOCATION: Mitochondrial membranes and perinuclear envelope (By similarity).

-!- DOMAIN: The BH4 domain is required for anti-apoptotic activity. The BH1 and BH2 domains are required for both heterodimerization with other Bcl-2 family members and for repression of cell death.

-!- PTM: Proteolytically cleaved by caspases during apoptosis (By similarity). The cleaved protein, lacking the BH4 domain, has pro-apoptotic activity (By similarity).

-!- SIMILARITY: Belongs to the Bcl-2 family.

-!- SIMILARITY: Contains 1 Bcl-2 homology 1 (BH1) domain.

-!- SIMILARITY: Contains 1 Bcl-2 homology 2 (BH2) domain.

-!- SIMILARITY: Contains 1 Bcl-2 homology 3 (BH3) domain.

-!- SIMILARITY: Contains 1 Bcl-2 homology 4 (BH4) domain.

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL; AJ001203; CAA04597.1; -

HSSP; Q07817; IMAZ.

InterPro; IPR000712; Bcl2 BH.

InterPro; IPR003093; Bcl2 BH4.

InterPro; IPR002475; BCL2 family.

InterPro; IPR004725; BCL2\_reg.

Pfam; PF00452; Bcl-2; 1.

Pfam; PF02180; BH4; 1.

TIGRFAMs; TIGR00865; bcl-2; 1.

PROSITE; PS00062; BCL2\_FAMILY; 1.

PROSITE; PS01080; BH1; 1.

PROSITE; PS01258; BH2; 1.

PROSITE; PS01259; BH3; 1.

PROSITE; PS01260; BH4; 1.

PROSITE; PS00063; BH4\_2; 1.

Apoptosis; Mitochondrion; Transmembrane.

DOMAIN 4 24 BH4.

DOMAIN 86 100 BH3.

DOMAIN 129 148 BH1.

DOMAIN 180 195 BH2.

TRANSMEM 210 226 Potential.

SEQUENCE 233 AA; 26061 MW; 18BF6FA0441912B2 CRC64;

Query Match 97.0%; Score 1181; DB 1; Length 233;  
Best Local Similarity 96.6%; Pred. No. 7e-94;  
Matches 225; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MSQSNRELVDVFLSYKLSQKGYSSQSDVENRTEAPEETPERETPSAINGNPSWHLA 60  
Db 1 MSQSNRELVDVFLSYKLSQKGYSSQSDVENRTEAPEETPERETPSAINGNPSWHLA 60

Qy 61 DSPAVNGATGHSSSLDAREVPMKALREAGDEFELRYRAFSDLTSLQHIPTGTAY 120  
Db 61 DSPAVNGATGHSSSLDAREVPMKALREAGDEFELRYRAFSDLTSLQHIPTGTAY 120

Qy 121 QSFQVNVNLFPRDGVNWRGRIVAFFSFGGALCVESVDKEMQVLVSRIASWMTATYLNHLEP 180  
Db 121 QSFQVNVNLFPRDGVNWRGRIVAFFSFGGALCVESVDKEMQVLVSRIASWMTATYLNHLEP 180

Qy 181 WIQENGWDVFDVLYGNNAASERKQGRFNRWFLTGMTVAGVLLGSLFSRK 233  
Db 181 WIQENGWDVFDVLYGNNAASERKQGRFNRWFLTGMTVAGVLLGSLFSRK 233

RESULT 8

Q8SQ42 ID Q8SQ42 PRELIMINARY; PRT; 233 AA.

AC Q8SQ42

DT 01-JUN-2002 (TrEMBLrel. 21, Created)

DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update).

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE Bcl-xl protein.

GN Name=bcl-xl;

OS Felis silvestris catus (Cat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.

OX NCBI\_TaxID=9685;

RN [1]

RP SEQUENCE FROM N.A.

RA Nagafuchi S., Sano J., Kano R., Hasegawa A.;

RL Submitted (MAR-2002) to the EMBL/GenBank/DBSJ databases.

DR EMBL; AB080951; BAB85856.2; -

DR HSSP; Q07817; IMAZ.

DR GO; GO:0016020; C:membrane; IEA.

DR GO; GO:0042981; P:regulation of apoptosis; IEA.

DR InterPro; IPR000712; Bcl2 BH.

DR InterPro; IPR003093; Bcl2 BH4.

DR InterPro; IPR002475; BCL2 family.

DR InterPro; IPR004725; BCL2\_reg.

Pfam; PF00452; Bcl-2; 1.

Pfam; PF02180; BH4; 1.

SMART; SM00337; BCL; 1.

SMART; SM00265; BH4; 1.

TIGRFAMs; TIGR00865; bcl-2; 1.

PROSITE; PS00062; BCL2\_FAMILY; 1.

PROSITE; PS01080; BH1; 1.

PROSITE; PS01258; BH2; 1.

PROSITE; PS01259; BH3; 1.

PROSITE; PS01260; BH4; 1.

PROSITE; PS00063; BH4\_2; 1.

SEQUENCE 233 AA; 26017 MW; CD17F24FE9D47BC9 CRC64;

Query Match 96.7%; Score 1178; DB 2; Length 233;  
Best Local Similarity 97.0%; Pred. No. 1.3e-93;  
Matches 226; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MSQSNRELVDVFLSYKLSQKGYSSQSDVENRTEAPEETPERETPSAINGNPSWHLA 60  
Db 1 MSQSNRELVDVFLSYKLSQKGYSSQSDVENRTEAPEETPERETPSAINGNPSWHLA 60

Qy 61 DSPAVNGATGHSSSLDAREVPMKALREAGDEFELRYRAFSDLTSLQHIPTGTAY 120  
Db 61 DSPAVNGATGHSSSLDAREVPMKALREAGDEFELRYRAFSDLTSLQHIPTGTAY 120

Qy 121 QSFQVNVNLFPRDGVNWRGRIVAFFSFGGALCVESVDKEMQVLVSRIASWMTATYLNHLEP 180  
Db 121 QSFQVNVNLFPRDGVNWRGRIVAFFSFGGALCVESVDKEMQVLVSRIASWMTATYLNHLEP 180

Qy 181 WIQENGWDVFDVLYGNNAASERKQGRFNRWFLTGMTVAGVLLGSLFSRK 233  
Db 181 WIQENGWDVFDVLYGNNAASERKQGRFNRWFLTGMTVAGVLLGSLFSRK 233

RESULT 9

Q9MYW4 ID Q9MYW4 PRELIMINARY; PRT; 233 AA.

AC Q9MYW4

DT 01-OCT-2000 (TrEMBLrel. 15, Created)

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE Bcl-X.

OS Oryctolagus cuniculus (Rabbit).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.

OX NCBI\_TaxID=9986;

RN [1]

RP SEQUENCE FROM N.A.

RA Knott J.C., Robertson L., James E.R.;

RL Submitted (JUL-2000) to the EMBL/GenBank/DBSJ databases.

DR EMBL; AY005131; AAF88137.1; -

DR HSSP; P53563; IAF3.

DR GO; GO:0016020; C:membrane; IEA.

DR GO; GO:0042981; P:regulation of apoptosis; IEA.  
 DR InterPro; IPR000712; Bcl2\_BH.  
 DR InterPro; IPR003093; Bcl2\_BH4.  
 DR InterPro; IPR002475; Bcl2\_family.  
 DR InterPro; IPR004725; Bcl2\_reg.  
 DR Pfam; PF00452; Bcl-2; 1.  
 DR Pfam; PF02180; BH4; 1.  
 DR SMART; SM00337; BCL; 1.  
 DR SMART; SM00265; BH4; 1.  
 DR TIGRFAMs; TIGR00865; bcl-2; 1.  
 DR PROSITE; PS00062; BCL2\_FAMILY; 1.  
 DR PROSITE; PS01080; BH1; 1.  
 DR PROSITE; PS01258; BH2; 1.  
 DR PROSITE; PS01259; BH3; 1.  
 DR PROSITE; PS01260; BH4; 1.  
 DR PROSITE; PS00063; BH4\_2; 1.  
 SQ SEQUENCE 233 AA; 25986 MW; 12F0F30344D53F93 CRC64;  
 Query Match 96.4%; Score 1174; DB 2; Length 233;  
 Best Local Similarity 96.1%; Pred. No. 2.8e-93;  
 Matches 224; Conservative 4; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 MSQSNRELVDVFLSYKLSQKGYSWQSDVEENRTEAPEETEPEETPSAINGNPSWHLA 60  
 DB 1 MSQSNRELVDVFLSYKLSQKGYSWQSDVEENRTEAPEETEPEETPSAINGNPSWHLA 60  
 QY 61 DSPAVNGATGHSSSLDAREVPMRAVKQALREAGDEPELRYRRAFSDLTSLQHTPGTAY 120  
 DB 61 DSPAVNGATGHSSSLDAREVPMRAVKQALREAGDEPELRYRRAFSDLTSLQHTPGTAY 120  
 QY 121 QSEFQVNVNLFDRDGNWGRIVAFPSFGGALCVESVDKEMQVLSRIASWMTYLNHLEP 180  
 DB 121 QSEFQVNVNLFDRDGNWGRIVAFPSFGGALCVESVDKEMQVLSRIASWMTYLNHLEP 180  
 QY 181 WIQENGWDTFTVDLYGNNAEESRKQGRFNRWFLTGMTVAGVLLGSLFSRK 233  
 DB 181 WIQENGWDTFTVDLYGNNAEESRKQGRFNRWFLTGMTVAGVLLGSLFSRK 233  
 RESULT 10  
 Q99N35 PRELIMINARY; PRT; 217 AA.  
 AC Q99N35; STRAIN=Sprague-Dawley;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE B-cell leukemia/lymphoma x (Fragment).  
 GN Name=Bclx;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129/SVJ;  
 RX MEDLINE=22202614; PubMed=12213327; DOI=10.1016/S0161-5890(02)00049-4;  
 RA Yang X.-F., Ye Q., Press B., Han R.-Z., Bassing C.H., Sleckman B.P.,  
 RA Alt F.W., Cantor H.;  
 RT "Analysis of the complex genomic structure of Bcl-x and its  
 RT relationship to Bcl-xgamma expression after CD28-dependent  
 RT costimulation."  
 RL Mol. Immunol. 39:45-55(2002).  
 DR EMBL; AF133282; AAK15455.1;  
 DR EMBL; AF133281; AAK15455.1; JOINED.  
 DR HSSP; P53563; IAF3.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0042981; P:regulation of apoptosis; IEA.  
 DR InterPro; IPR000712; Bcl2\_BH.  
 DR InterPro; IPR002475; Bcl2\_family.  
 DR InterPro; IPR004725; Bcl2\_reg.  
 DR Pfam; PF00452; Bcl-2; 1.  
 DR SMART; SM00337; BCL; 1.  
 DR TIGRFAMs; TIGR00865; bcl-2; 1.

DR PROSITE; PS00062; BCL2\_FAMILY; 1.  
 DR PROSITE; PS01080; BH1; 1.  
 DR PROSITE; PS01258; BH2; 1.  
 DR PROSITE; PS01259; BH3; 1.  
 FT NON\_TER 1  
 SQ SEQUENCE 217 AA; 24234 MW; 3B5A4E809A7DEF18 CRC64;  
 Query Match 92.9%; Score 1132; DB 2; Length 217;  
 Best Local Similarity 99.5%; Pred. No. 1.1e-89;  
 Matches 216; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 17 LSQKGYSWQSDVEENRTEAPEETEPEETPSAINGNPSWHLADSPAVNGATGHSSSLD 76  
 DB 1 LSQKGYSWQSDVEENRTEAPEETEPEETPSAINGNPSWHLADSPAVNGATGHSSSLD 60  
 QY 77 AREVPMRAVKQALREAGDEPELRYRRAFSDLTSLQHTPGTAYQSEFQVNVNLFDRDGN 136  
 DB 61 AREVPMRAVKQALREAGDEPELRYRRAFSDLTSLQHTPGTAYQSEFQVNVNLFDRDGN 120  
 QY 137 WGRIVAFPSFGGALCVESVDKEMQVLSRIASWMTYLNHLEPWIQENGWDTFTVDLYG 196  
 DB 121 WGRIVAFPSFGGALCVESVDKEMQVLSRIASWMTYLNHLEPWIQENGWDTFTVDLYG 180  
 QY 197 NNAEESRKQGRFNRWFLTGMTVAGVLLGSLFSRK 233  
 DB 181 NNAEESRKQGRFNRWFLTGMTVAGVLLGSLFSRK 217  
 RESULT 11  
 Q7TS62 PRELIMINARY; PRT; 284 AA.  
 AC Q7TS62; STRAIN=Sprague-Dawley;  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Bcl-xbeta.  
 GN Name=Bcl-x;  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Sprague-Dawley;  
 RX MEDLINE=22672518; PubMed=12787069;  
 RA Itoh T., Itoh A., Pleasure D.;  
 RT "Bcl-2-related protein family gene expression during oligodendroglial  
 RT differentiation."  
 RL J. Neurochem. 85:1500-1512(2003).  
 DR EMBL; AY141038; AAN17784.1; -.  
 DR HSSP; P53563; IAF3.  
 DR GO; GO:0042981; P:regulation of apoptosis; IEA.  
 DR InterPro; IPR000712; Bcl2\_BH.  
 DR InterPro; IPR003093; Bcl2\_BH4.  
 DR InterPro; IPR002475; BCL2\_family.  
 DR Pfam; PF00452; Bcl-2; 1.  
 DR Pfam; PF02180; BH4; 1.  
 DR SMART; SM00337; BCL; 1.  
 DR SMART; SM00265; BH4; 1.  
 DR PROSITE; PS00062; BCL2\_FAMILY; 1.  
 DR PROSITE; PS01080; BH1; 1.  
 DR PROSITE; PS01259; BH3; 1.  
 DR PROSITE; PS01260; BH4; 1.  
 DR PROSITE; PS00063; BH4\_2; 1.  
 SQ SEQUENCE 284 AA; 31776 MW; B8F35F641D4E029E CRC64;  
 Query Match 81.0%; Score 987; DB 2; Length 284;  
 Best Local Similarity 100.0%; Pred. No. 5.3e-77;  
 Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MSQSNRELVDVFLSYKLSQKGYSWQSDVEENRTEAPEETEPEETPSAINGNPSWHLA 60  
 DB 1 MSQSNRELVDVFLSYKLSQKGYSWQSDVEENRTEAPEETEPEETPSAINGNPSWHLA 60



Qy 61 DSPAVNGATGSSSLDAREVTPMAAVKQALREAGDEFELRYRRAFSDLTSQLHTTPGTAY 120  
 |||||  
 Db 61 DSPAVNGATGSSSLDAREVTPMAAVKQALREAGDEFELRYRRAFSDLTSQLHTTPGTAY 120  
 |||||  
 Qy 121 QSFQVNVNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLVSRIASWMTATYLNHLEP 180  
 |||||  
 Db 121 QSFQVNVNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLVSRIASWMTATYLNHLEP 180  
 |||||  
 Qy 181 WIQENGW 188  
 |||||  
 Db 181 WIQENGW 188  
 |||||  
 RESULT 12  
 Q9QWX2 PRELIMINARY; PRT; 188 AA.  
 AC Q9QWX2;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Bcl-x (Fragment).  
 GN Name=Bcl2l1; Synonyms=Bcl2l;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCB1\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=20350651; PubMed=10894153; DOI=10.1210/me.14.7.1038;  
 RX Rucker E.B. III, Dierisseau P., Wagner K.U., Garrett L.,  
 RA Wynshaw-Boris A., Flaws J.A., Hennighausen L.;  
 RT "Bcl-x and Bax regulate mouse primordial germ cell survival and  
 apoptosis during embryogenesis.";  
 RL Mol. Endocrinol. 14:1038-1052(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Rucker E. III, Dierisseau P., Herring S., Wagner K.-U.,  
 RA Hennighausen L.;  
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF088904; AAC7232.1; --  
 DR HSSP; P53563; 1AF3.  
 DR MGD; MGI:88139; Bcl2l1.  
 DR GO; GO:0005739; C:mitochondrion; IDA.  
 DR GO; GO:0005515; F:protein binding; IPI.  
 DR GO; GO:0006916; P:anti-apoptosis; IDA.  
 DR GO; GO:0006915; P:apoptosis; IDA.  
 DR GO; GO:0009314; P:response to radiation; IMP.  
 DR InterPro; IPR000712; Bcl2 BH.  
 DR InterPro; IPR003093; Bcl2 family.  
 DR Pfam; PF00452; Bcl-2; 1.  
 DR Pfam; PF02180; BH4; 1.  
 DR SMART; SM00337; BCL; 1.  
 DR SMART; SM00265; BH4; 1.  
 DR TIGRFAMs; TIGR00865; bcl-2; 1.  
 DR PROSITE; PS00062; BCL2\_FAMILY; 1.  
 DR PROSITE; PS01080; BH1; 1.  
 DR PROSITE; PS01259; BH3; 1.  
 DR PROSITE; PS01260; BH4; 1.  
 DR PROSITE; PS00063; BH4\_2; 1.  
 FT NON TER 188  
 SQ SEQUENCE 188 AA; 21126 MW; 4562F8356D248E52 CRC64;  
 Query Match 80.4%; Score 979; DB 2; Length 188;  
 Best Local Similarity 99.5%; Pred. No. 1.5e-76;  
 Matches 187; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 MSQSNRELVDVFLSYKLSQKGYSWQSFSDVEENRTEAPEETEPETPSAINGNPSWHLA 60  
 |||||  
 Db 1 MSQSNRELVDVFLSYKLSQKGYSWQSFSDVEENRTEAPEETEPETPSAINGNPSWHLA 60  
 |||||  
 Qy 61 DSPAVNGATGSSSLDAREVTPMAAVKQALREAGDEFELRYRRAFSDLTSQLHTTPGTAY 120  
 |||||

Db 61 DSPAVNGATGSSSLDAREVTPMAAVKQALREAGDEFELRYRRAFSDLTSQLHTTPGTAY 120  
 |||||  
 Qy 121 QSFQVNVNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLVSRIASWMTATYLNHLEP 180  
 |||||  
 Db 121 QSFQVNVNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLVSRIASWMTATYLNHLEP 180  
 |||||  
 Qy 181 WIQENGW 188  
 |||||  
 Db 181 WIQENGW 188  
 |||||  
 RESULT 13  
 Q35843 PRELIMINARY; PRT; 235 AA.  
 AC Q35843;  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Bcl-x-gamma.  
 GN Name=Bcl2l1; Synonyms=Bcl2l;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCB1\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=B6/CBA; TISSUE=Thymus;  
 RX MEDLINE=98051053; PubMed=9390687; DOI=10.1016/S1074-7613(00)80384-2;  
 RA Yang X.-F., Weber G.F., Cantor H.;  
 RT "A novel Bcl-x isoform connected to the T cell receptor regulates  
 apoptosis in T cells.";  
 RL Immunity 7:629-639(1997).  
 DR EMBL; U51277; AAC53458.1; --  
 DR HSSP; P53563; 1AF3.  
 DR MGD; MGI:88139; Bcl2l1.  
 DR GO; GO:0005739; C:mitochondrion; IDA.  
 DR GO; GO:0005515; F:protein binding; IPI.  
 DR GO; GO:0006916; P:anti-apoptosis; IDA.  
 DR GO; GO:0006915; P:apoptosis; IDA.  
 DR GO; GO:0009314; P:response to radiation; IMP.  
 DR InterPro; IPR000712; Bcl2 BH.  
 DR InterPro; IPR003093; Bcl2 family.  
 DR InterPro; IPR002475; Bcl2\_family.  
 DR Pfam; PF00452; Bcl-2; 1.  
 DR Pfam; PF02180; BH4; 1.  
 DR SMART; SM00337; BCL; 1.  
 DR SMART; SM00265; BH4; 1.  
 DR TIGRFAMs; TIGR00865; bcl-2; 1.  
 DR PROSITE; PS00062; BCL2\_FAMILY; 1.  
 DR PROSITE; PS01080; BH1; 1.  
 DR PROSITE; PS01259; BH3; 1.  
 DR PROSITE; PS01260; BH4; 1.  
 DR PROSITE; PS00063; BH4\_2; 1.  
 FT NON TER 188  
 SQ SEQUENCE 235 AA; 26122 MW; 649D914C2D5378F6 CRC64;  
 Query Match 80.4%; Score 979; DB 2; Length 235;  
 Best Local Similarity 99.5%; Pred. No. 2e-76;  
 Matches 187; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 MSQSNRELVDVFLSYKLSQKGYSWQSFSDVEENRTEAPEETEPETPSAINGNPSWHLA 60  
 |||||  
 Db 1 MSQSNRELVDVFLSYKLSQKGYSWQSFSDVEENRTEAPEETEPETPSAINGNPSWHLA 60  
 |||||  
 Qy 61 DSPAVNGATGSSSLDAREVTPMAAVKQALREAGDEFELRYRRAFSDLTSQLHTTPGTAY 120  
 |||||  
 Db 61 DSPAVNGATGSSSLDAREVTPMAAVKQALREAGDEFELRYRRAFSDLTSQLHTTPGTAY 120  
 |||||  
 Qy 121 QSFQVNVNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLVSRIASWMTATYLNHLEP 180  
 |||||  
 Db 121 QSFQVNVNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLVSRIASWMTATYLNHLEP 180  
 |||||  
 Qy 181 WIQENGW 188  
 |||||



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2005, 00:22:11 ; Search time 30.4692 Seconds  
(without alignments)  
735.775 Million cell updates/sec

Title: US-10-049-822A-2  
Perfect score: 1218  
Sequence: 1 MSQSNRELVDVFLSYKLSQK.....FLGTMVAGVVLGSLFSRK 233

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1214	99.7	233	2 S51761	BCL-X protein - ra
2	1210	99.3	233	2 I49056	bcl-x long - mouse
3	1187	97.5	233	2 I67431	BCL-X-Long - rat
4	1184	97.2	233	2 B47537	apoptosis regulato
5	1006	82.6	214	2 I49057	bcl-x transmembran
6	963	79.1	227	2 JB0203	apoptosis regulato
7	825.5	67.8	170	2 I49055	bcl-x short - mous
8	818.5	67.2	176	2 I67435	gene bcl-xshort pr
9	732	60.1	190	2 A47537	apoptosis regulato
10	497	40.8	233	2 A37332	transforming prote
11	474.5	39.0	232	2 S24390	transforming prote
12	465.5	38.2	236	2 JG7383	B-cell lymphoma 2
13	465	38.2	236	2 I67432	BCL-2 - rat (fragm
14	464	38.1	239	1 TVHU11	transforming prote
15	460.5	37.8	236	2 I53744	gene bcl-2 protein
16	457	37.5	236	1 TVWSA1	transforming prote
17	447.5	36.7	216	2 B37332	transforming prote
18	423.5	34.8	205	1 TVHU31	transforming prote
19	422	34.6	199	1 TVWSB1	transforming prote
20	315	25.9	154	2 I58194	gene bcl-2 protein
21	182	14.9	211	2 S58875	cdn-2 protein - hu
22	181	14.9	211	2 S58873	Bak protein - huma
23	167	13.7	177	2 S54778	NR-13 protein - qu
24	166	13.6	192	2 D47538	bcl-2-associated p
25	166	13.6	192	2 A47538	bcl-2-associated p
26	165.5	13.6	133	2 I53295	bcl-2-associated p
27	152	12.5	143	2 I38921	bcl-2-associated p
28	150.5	12.4	179	2 JC7255	Bax-delta protein
29	148.5	12.2	218	2 B47538	bcl-2-associated p

30	148	12.2	255	2 JC7567	Mcl-1a protein - 2
31	144	11.8	172	2 I49449	hemopoietic-specif
32	137.5	11.3	350	2 A47476	BCL2 homolog MCL1
33	135	11.1	261	2 H88578	protein ced-9 (imp
34	135	11.1	280	2 A53189	apoptosis suppress
35	130	10.7	175	2 I39055	Bcl-2 related - hu
36	102	8.4	160	2 G36807	hypothetical prote
37	94.5	7.8	191	1 Q0B54	BHRF1 protein - hu
38	93	7.6	780	2 T09485	cold-induced prote
39	91	7.5	801	2 B83195	hypothetical prote
40	91	7.5	1215	2 E84480	probable retroelem
41	89.5	7.3	462	2 E82683	RNA polymerase sig
42	87	7.1	442	2 T39883	zucotin-like protei
43	86.5	7.1	1061	1 DJAD12	DNA-directed DNA p
44	86	7.1	544	2 S19149	thioglucoisidase (E
45	86	7.1	658	2 S37494	squalene-hopene cy

ALIGNMENTS

RESULT 1

S51761  
BCL-X protein - rat  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 07-May-1995 #sequence\_revision 01-Sep-1995 #text\_change 09-Jul-2004  
C;Accession: S51761; S51762  
R;Michaelidis, T.M.  
submitted to the EMBL Data Library, November 1994  
A;Reference number: S51761  
A;Accession: S51761  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-233 <MIC>  
A;Cross-references: UNIPROT:P53563; EMBL:X82537; NID:g607176; PIDN:CAA57886.1; PID:g6071  
A;Experimental source: embryonic; brain  
A;Accession: S51762  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-125,189-233 <MI2>  
A;Cross-references: EMBL:X82537; NID:g607176; PIDN:CAA57887.1; PID:g607178  
A;Experimental source: embryonic; brain  
A;Note: smaller form due to splicing  
C;Genetics:  
A;Introns: 125/3  
C;Superfamily: bcl apoptosis regulator, inhibitory type

Query Match	99.7%;	Score 1214;	DB 2;	Length 233;
Best Local Similarity	99.6%;	Pred. No. 9.6e-99;		
Matches	232;	Conservative	1;	Mismatches 0; Indels 0; Gaps 0;
Qy	1	MSQSNRELVDVFLSYKLSQKYSWSQFSQSDVEENRTEAPEETEPEETPSAINGNPSWHLA	60	
Db	1	MSQSNQSLVDVFLSYKLSQKYSWSQFSQSDVEENRTEAPEETEPEETPSAINGNPSWHLA	60	
Qy	61	DSPAVNGATGCHSSSLDAREVPMAAVKQALREAGDEFELRYRRAFSDLTSLHTTPGTAY	120	
Db	61	DSPAVNGATGCHSSSLDAREVPMAAVKQALREAGDEFELRYRRAFSDLTSLHTTPGTAY	120	
Qy	121	QSPFQVNNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLVSRIASWMTATYNDHLEP	180	
Db	121	QSPFQVNNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLVSRIASWMTATYNDHLEP	180	
Qy	181	WIQENGGWDTFVDLYGNNAAESKQGRFNRNFWLTGTMVAGVVLGSLFSRK	233	
Db	181	WIQENGGWDTFVDLYGNNAAESKQGRFNRNFWLTGTMVAGVVLGSLFSRK	233	

RESULT 2

I49056  
bcl-x long - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 09-Jul-2004

C;Accession: I49056; S52866  
R;Fang, W.; Rivard, J.J.; Mueller, D.L.; Behrens, T.W.  
J. Immunol. 153, 4388-4398, 1994  
A;Title: Cloning and molecular characterization of mouse bcl-x in B and T lymphocytes.  
A;Reference number: I49055; MUID:95052604; PMID:7963517  
A;Accession: I49056  
A;Status: preliminary; translated from GB/EMBL/DBDJ  
A;Molecule type: mRNA  
A;Residues: 1-233 <RES>  
A;Cross-references: UNIPROT:Q64373; EMBL:U10101; NID:G506647; PIDN:AAA82173.1; PID:G5066  
R;Kamesaki, H.; Michaud, G.Y.; Takatsu, K.; Okuma, M.  
submitted to the EMBL Data Library, November 1994  
A;Description: IL-5 inhibits anti-IgM-induced apoptosis in an immature B cell line through  
A;Reference number: S52866  
A;Accession: S52866  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-233 <KAW>  
A;Cross-references: EMBL:X83574; NID:G695622; PIDN:CAA58557.1; PID:G695623  
C;Superfamily: bcl apoptosis regulator, inhibitory type

Query Match 99.3%; Score 1210; DB 2; Length 233;  
Best Local Similarity 99.6%; Pred. No. 2.1e-98;  
Matches 232; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MSQSNRELVDVFLSYKLSQKGYSWQSFSDVEENRTEAPEETEPEETPSAINGNPSWHLA 60  
Db 1 MSQSNRELVDVFLSYKLSQKGYSWQSFSDVEENRTEAPEETEPEETPSAINGNPSWHLA 60

Qy 61 DSPAVNGATGHSSSLDAREVIPMAAVKQALREAGDEFEFLRYRRAFSDLTSQLHTPGTAY 120  
Db 61 DSPAVNGATGHSSSLDAREVIPMAAVKQALREAGDEFEFLRYRRAFSDLTSQLHTPGTAY 120

Qy 121 QSFQVNVNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLVSRIASWMATYLNHLEP 180  
Db 121 QSFQVNVNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLVSRIASWMATYLNHLEP 180

Qy 181 WIQENGWDTFVDLYGNNAAESRKQGERFNRWFLTGMTVAGVLLGSLFSRK 233  
Db 181 WIQENGWDTFVDLYGNNAAESRKQGERFNRWFLTGMTVAGVLLGSLFSRK 233

RESULT 3  
BCL2-X-Long - rat  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 26-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 09-Jul-2004  
C;Accession: I67431  
R;Tilly, J.L.; Tilly, K.I.; Kenton, M.L.; Johnson, A.L.  
Endocrinology 136, 232-241, 1995  
A;Title: Expression of members of the bcl-2 gene family in the immature rat ovary: equivalent  
constitutive bcl-2 and bcl-x long messenger ribonucleic acid levels.  
A;Reference number: I53295; MUID:95129487; PMID:7828536  
A;Accession: I67431  
A;Status: preliminary; translated from GB/EMBL/DBDJ  
A;Molecule type: mRNA  
A;Residues: 1-233 <RES>  
A;Cross-references: UNIPROT:P53563; EMBL:U34963; NID:G1004376; PIDN:AAA77686.1; PID:G100  
C;Superfamily: bcl apoptosis regulator, inhibitory type

Query Match 97.5%; Score 1187; DB 2; Length 233;  
Best Local Similarity 97.4%; Pred. No. 2.2e-96;  
Matches 227; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 MSQSNRELVDVFLSYKLSQKGYSWQSFSDVEENRTEAPEETEPEETPSAINGNPSWHLA 60  
Db 1 MSQSNRELVDVFLSYKLSQKGYSWQSFSDVEENRTEAPEETEPEETPSAINGNPSWHLA 60

Qy 61 DSPAVNGATGHSSSLDAREVIPMAAVKQALREAGDEFEFLRYRRAFSDLTSQLHTPGTAY 120  
Db 61 DSPAVNGATGHSSSLDAREVIPMAAVKQALREAGDEFEFLRYRRAFSDLTSQLHTPGTAY 120

Qy 121 QSFQVNVNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLVSRIASWMATYLNHLEP 180  
Db 121 QSFQVNVNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLVSRIASWMATYLNHLEP 180

Qy 121 QSFQVNVNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLVSRIASWMATYLNHLEP 180

Db 121 QSFQVNVNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLVSRIASWMATYLNHLEP 180

Qy 181 WIQENGWDTFVDLYGNNAAESRKQGERFNRWFLTGMTVAGVLLGSLFSRK 233  
Db 181 WIQENGWDTFVDLYGNNAAESRKQGERFNRWFLTGMTVAGVLLGSLFSRK 233

RESULT 4  
apoptosis regulator bcl-xL - human  
N;Alternate names: bcl-2-related protein  
N;Contains: apoptosis regulator bcl-xs  
C;Species: Homo sapiens (man)  
C;Date: 16-Aug-1996 #sequence\_revision 16-Aug-1996 #text\_change 09-Jul-2004  
C;Accession: B47537  
R;Boise, L.H.; Gonzalez-Garcia, M.; Postema, C.E.; Ding, L.; Lindsten, T.; Turka, L.A.;  
Cell 74, 597-608, 1993  
A;Title: bcl-x, a bcl-2-related gene that functions as a dominant regulator of apoptotic  
A;Reference number: A47537; MUID:93364977; PMID:8358789  
A;Accession: B47537  
A;Status: nucleic acid sequence not shown; translated from GB/EMBL/DBDJ  
A;Molecule type: mRNA  
A;Residues: 1-233 <BOI>  
A;Cross-references: UNIPROT:Q07817; GB:L20121; NID:G510900; PIDN:CAA80661.1; PID:G510901  
A;Accession: C47537  
A;Status: nucleic acid sequence not shown; translated from GB/EMBL/DBDJ  
A;Molecule type: mRNA  
A;Residues: 1-69, 'G', 71-125, 189-233 <BO2>  
A;Cross-references: GB:L20122; NID:G623236; PIDN:CAA80662.1; PID:G623237  
C;Genetics:  
A;Gene: GDB:BCL2L  
A;Cross-references: GDB:228079  
C;Superfamily: bcl apoptosis regulator, inhibitory type  
C;Keywords: alternative splicing; apoptosis  
F;1-233/Product: apoptosis regulator bcl-xL #status predicted <MAT>  
F;1-125,189-233/Product: apoptosis regulator bcl-xs #status predicted <MA2>

Query Match 97.2%; Score 1184; DB 2; Length 233;  
Best Local Similarity 97.4%; Pred. No. 4e-96;  
Matches 227; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MSQSNRELVDVFLSYKLSQKGYSWQSFSDVEENRTEAPEETEPEETPSAINGNPSWHLA 60  
Db 1 MSQSNRELVDVFLSYKLSQKGYSWQSFSDVEENRTEAPEETEPEETPSAINGNPSWHLA 60

Qy 61 DSPAVNGATGHSSSLDAREVIPMAAVKQALREAGDEFEFLRYRRAFSDLTSQLHTPGTAY 120  
Db 61 DSPAVNGATGHSSSLDAREVIPMAAVKQALREAGDEFEFLRYRRAFSDLTSQLHTPGTAY 120

Qy 121 QSFQVNVNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLVSRIASWMATYLNHLEP 180  
Db 121 QSFQVNVNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLVSRIASWMATYLNHLEP 180

Qy 181 WIQENGWDTFVDLYGNNAAESRKQGERFNRWFLTGMTVAGVLLGSLFSRK 233  
Db 181 WIQENGWDTFVDLYGNNAAESRKQGERFNRWFLTGMTVAGVLLGSLFSRK 233

RESULT 5  
bcl-x transmembrane deleted - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 09-Jul-2004  
C;Accession: I49057  
R;Fang, W.; Rivard, J.J.; Mueller, D.L.; Behrens, T.W.  
J. Immunol. 153, 4388-4398, 1994  
A;Title: Cloning and molecular characterization of mouse bcl-x in B and T lymphocytes.  
A;Reference number: I49055; MUID:95052604; PMID:7963517  
A;Accession: I49057  
A;Status: preliminary; translated from GB/EMBL/DBDJ  
A;Molecule type: mRNA  
A;Residues: 1-214 <RES>

A;Cross-references: UNIPROT:Q64373; EMBL:U10102; NID:G506649; PIDN:AAA82174.1; PID:G50664  
 C;Genetics:  
 A;Gene: bcl-x-long  
 C;Superfamily: bcl apoptosis regulator, inhibitory type

Query Match 82.6%; Score 1006; DB 2; Length 214;  
 Best Local Similarity 99.5%; Pred. No. 1.3e-80;  
 Matches 192; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MSQSNRELVDVFLSYKLSQKGYSSQSDVEENRTEAPEETEPEERETPSAINGNPSWHLA 60  
 Db 1 MSQSNRELVDVFLSYKLSQKGYSSQSDVEENRTEAPEETEPEERETPSAINGNPSWHLA 60

Qy 61 DSPAVNGATGHSSSLDAREVIPMAAVKQALREAGDEPELRYRRAFSDLTSQLHTTPGTAY 120  
 Db 61 DSPAVNGATGHSSSLDAREVIPMAAVKQALREAGDEPELRYRRAFSDLTSQLHTTPGTAY 120

Qy 121 QSFQVNVNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLVSRIASWMTATYLNHLEP 180  
 Db 121 QSFQVNVNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLVSRIASWMTATYLNHLEP 180

Qy 181 WIQENGWDTFVD 193  
 Db 181 WIQENGWDTFVD 193

RESULT 6  
 JE0203  
 apoptosis regulator bcl-x isoform - human  
 N;Alternate names: h-bcl-xbeta  
 C;Species: Homo sapiens (man)  
 C;Date: 21-Aug-1998 #sequence\_revision 21-Aug-1998 #text\_change 09-Jul-2004  
 C;Accession: JE0203  
 R;Ban, J.; Eckhart, L.; Weninger, W.; Mildner, M.; Tschachler, E.  
 Biochem. Biophys. Res. Commun. 248, 147-152, 1998  
 A;Title: Identification of a human cDNA encoding a novel bcl-x isoform.  
 A;Reference number: JE0203; MUID:98340865; PMID:9675101  
 A;Accession: JE0203  
 A;Molecule type: mRNA  
 A;Residues: 1-227 <BAN>  
 A;Cross-references: UNIPROT:Q07817; GB:U72398; NID:G1622940; PIDN:AA817354.1; PID:G16229  
 C;Genetics:  
 A;Gene: bcl-x  
 A;Map position: 20  
 C;Superfamily: bcl apoptosis regulator, inhibitory type

Query Match 79.1%; Score 963; DB 2; Length 227;  
 Best Local Similarity 97.9%; Pred. No. 8.3e-77;  
 Matches 184; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MSQSNRELVDVFLSYKLSQKGYSSQSDVEENRTEAPEETEPEERETPSAINGNPSWHLA 60  
 Db 1 MSQSNRELVDVFLSYKLSQKGYSSQSDVEENRTEAPEETEPEERETPSAINGNPSWHLA 60

Qy 61 DSPAVNGATGHSSSLDAREVIPMAAVKQALREAGDEPELRYRRAFSDLTSQLHTTPGTAY 120  
 Db 61 DSPAVNGATGHSSSLDAREVIPMAAVKQALREAGDEPELRYRRAFSDLTSQLHTTPGTAY 120

Qy 121 QSFQVNVNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLVSRIASWMTATYLNHLEP 180  
 Db 121 QSFQVNVNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLVSRIASWMTATYLNHLEP 180

Qy 181 WIQENGW 188  
 Db 181 WIQENGW 188

RESULT 7  
 I49055  
 bcl-x short - mouse  
 C;Species: Mus musculus (house mouse)  
 C;Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 09-Jul-2004  
 C;Accession: I49055

R;Fang, W.; Rivard, J.J.; Mueller, D.L.; Behrens, T.W.  
 J. Immunol. 153, 4388-4398, 1994  
 A;Title: Cloning and molecular characterization of mouse bcl-x in B and T lymphocytes.  
 A;Reference number: I49055; MUID:95052604; PMID:7963517  
 A;Accession: I49055  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: mRNA  
 A;Residues: 1-170 <RES>  
 A;Cross-references: UNIPROT:Q64373; EMBL:U10100; NID:G506645; PIDN:AAA82172.1; PID:G50664  
 C;Genetics:  
 A;Gene: bcl-x  
 C;Superfamily: bcl apoptosis regulator, inhibitory type

Query Match 67.8%; Score 825.5; DB 2; Length 170;  
 Best Local Similarity 72.5%; Pred. No. 6e-65;  
 Matches 169; Conservative 0; Mismatches 1; Indels 63; Gaps 1;

Qy 1 MSQSNRELVDVFLSYKLSQKGYSSQSDVEENRTEAPEETEPEERETPSAINGNPSWHLA 60  
 Db 1 MSQSNRELVDVFLSYKLSQKGYSSQSDVEENRTEAPEETEPEERETPSAINGNPSWHLA 60

Qy 61 DSPAVNGATGHSSSLDAREVIPMAAVKQALREAGDEPELRYRRAFSDLTSQLHTTPGTAY 120  
 Db 61 DSPAVNGATGHSSSLDAREVIPMAAVKQALREAGDEPELRYRRAFSDLTSQLHTTPGTAY 120

Qy 121 QSFQVNVNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLVSRIASWMTATYLNHLEP 180  
 Db 121 QSFQVNVNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLVSRIASWMTATYLNHLEP 180

Qy 181 WIQENGWDTFVDLYGNNAASERKGOERNRWFELTGMTVAGVVLGSLFSRK 233  
 Db 126 -----DTFVDLYGNNAASERKGOERNRWFELTGMTVAGVVLGSLFSRK 170

RESULT 8  
 I67435  
 gene bcl-x short protein - rat (fragment)  
 C;Species: Rattus sp. (rat)  
 C;Date: 26-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 28-Jul-2003  
 C;Accession: I67435  
 R;Tilly, J.L.; Tilly, K.I.; Kenton, M.L.; Johnson, A.L.  
 Endocrinology 136, 232-241, 1995  
 A;Title: Expression of members of the bcl-2 gene family in the immature rat ovary: equin  
 onstitutive bcl-2 and bcl-x long messenger ribonucleic acid levels.  
 A;Reference number: I53295; MUID:95129487; PMID:7828536  
 A;Accession: I67435  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: mRNA  
 A;Residues: 1-176 <RES>  
 A;Cross-references: GB:S78284; NID:G998483; PIDN:AA60702.1; PID:G998484  
 C;Genetics:  
 A;Gene: bcl-x  
 C;Superfamily: bcl apoptosis regulator, inhibitory type

Query Match 67.2%; Score 818.5; DB 2; Length 176;  
 Best Local Similarity 71.2%; Pred. No. 2.6e-64;  
 Matches 166; Conservative 1; Mismatches 3; Indels 63; Gaps 1;

Qy 1 MSQSNRELVDVFLSYKLSQKGYSSQSDVEENRTEAPEETEPEERETPSAINGNPSWHLA 60  
 Db 7 MSQSNRELVDVFLSYKLSQKGYSSQSDVEENRTEAPEETEPEERETPSAINGNPSWHLA 66

Qy 61 DSPAVNGATGHSSSLDAREVIPMAAVKQALREAGDEPELRYRRAFSDLTSQLHTTPGTAY 120  
 Db 67 DSPAVNGATGHSSSLDAREVIPMAAVKQALREAGDEPELRYRRAFSDLTSQLHTTPGTAY 126

Qy 121 QSFQVNVNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLVSRIASWMTATYLNHLEP 180  
 Db 127 QSFQVNVNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLVSRIASWMTATYLNHLEP 180

Qy 181 WIQENGWDTFVDLYGNNAASERKGOERNRWFELTGMTVAGVVLGSLFSRK 233  
 Db 132 -----DTFVDLYGNNTAPEERKGOERNRWFELTGMTVAGVVLGSLFSRK 176



Db 11 NREIWMYIHYKLSQRYEW-DVGDVD-----AAPLGAAP---TFGIFSFQPSN--PTPA 60  
Qy 65 VN-----GATGSSSLDAREVPMVAQALREAGDEPELRYRRFAFSDLTLS 110  
Db 61 VHRDMAARTSLPRIVATTGPTLS-----PVPVPV--VHLTLRAGDDFSRRYRDFAEWSS 114  
Qy 111 QLHITPGTAYQSFQVNNELFRDGVNMGRIIVAFPSFGALCVESVDKEMQVLSVRIASWM 170  
Db 115 QLHLTPPTARGFATVVEELFRDGVNMGRIIVAFPEFGVCMCVESVNRMSPLVDNIALWM 174  
Qy 171 ATYLNHLEPWIQNGGWDTFVDLYGNNAEESRKQERFNRWELTGMTVAGVVLGS 228  
Db 175 TEYLNRHLHTWIQNGGWDFAVELYGPSV-----RPLDFSWLSLTKLLSLALVGA 225

RESULT 13  
167432  
BCL-2 - rat (fragment)  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 26-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 09-Jul-2004  
C:Accession: I67432  
R:Tilly, J.L.; Tilly, K.I.; Kenton, M.L.; Johnson, A.L.  
Endocrinology 136, 232-241, 1995  
A:Title: Expression of members of the bcl-2 gene family in the immature rat ovary: equin  
onstitutive bcl-2 and bcl-2-xiong messenger ribonucleic acid levels.  
A:Reference number: I53295; MUID:95129487; PMID:7828536  
A:Accession: I67432  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-236 <RES>  
A:Cross-references: UNIPROT:P49950; EMBL:U34964; NID:g1004378; PIDN:AAA7687.1; PID:g100  
C:Superfamily: bcl apoptosis regulator, inhibitory type

Query Match 38.2%; Score 465; DB 2; Length 236;  
Best Local Similarity 48.3%; Pred. No. 3.2e-33;  
Matches 98; Conservative 25; Mismatches 56; Indels 24; Gaps 5;

Qy 5 NREIWDPLSYKLSQKYSWSQFSDVENRTEAPEETEPEPE--TPSAINGNPSWHLADS 62  
Db 11 NREIWMYIHYKLSQRYEW-D-----TGDEDSAPLRAPTGGIFSFQPSN--RT 58  
Qy 63 PAVNGATGSHSSLDAREVPMVA-----AVQALREAGDEPELRYRRFAFSDLTSLH 113  
Db 59 PAVHRDTAARTS-PLRPLVANAGPALSPPVPVHLLTLRAGDDFSRRYRDFAEWSSQLH 117  
Qy 114 ITPGTAYQSFQVNNELFRDGVNMGRIIVAFPSFGALCVESVDKEMQVLSVRIASWMATY 173  
Db 118 LTPPTARGFATVVEELFRDGVNMGRIIVAFPEFGVCMCVESVNRMSPLVDNIALWMTEY 177  
Qy 174 LNDHLEPWIQNGGWDTFVDLYG 196  
Db 178 LNRHLHTWIQNGGWDFAVELYG 200

RESULT 14  
TVHUA1  
transforming protein bcl-2, splice form alpha - human  
C:Species: Homo sapiens (man)  
C:Date: 31-Dec-1988 #sequence\_revision 07-Jun-1996 #text\_change 09-Jul-2004  
C:Accession: C37332; A29409; S02452; A24428; A27622; B27622  
R:Eguchi, Y.; Ewert, D.L.; Tsujimoto, Y.  
Nucleic Acids Res. 20, 4187-4192, 1992  
A:Title: Isolation and characterization of the chicken bcl-2 gene: expression in a varie  
A:Reference number: A37332; MUID:92375724; PMID:1508712  
A:Accession: C37332  
A:Status: nucleic acid sequence not shown; not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-239 <EGU>  
A:Cross-references: UNIPROT:P10415  
A:Note: this report is a correction  
R:Tsujimoto, Y.; Croce, C.M.  
Proc. Natl. Acad. Sci. U.S.A. 83, 5214-5218, 1986

A:Title: Analysis of the structure, transcripts, and protein products of bcl-2, the gene  
A:Reference number: A29409; MUID:86259760; PMID:3523487  
A:Accession: A29409  
A:Molecule type: mRNA  
A:Residues: 1-95, 'A', 'G', '111-236, 'S', '238-239 <TSU>  
A:Cross-references: GB:M1394; NID:g179366; PIDN:AAA51813.1; PID:g179367  
A:Note: this sequence has been corrected in reference A37332  
R:Sato, M.; Jaeger, U.; Hockett, R.D.; Graninger, W.; Bennett, S.; Goldman, P.; Korsmeyer  
EMBO J. 7, 123-131, 1988  
A:Title: Alternative promoters and exons, somatic mutation and deregulation of the Bcl-2  
A:Reference number: S02452; MUID:88196071; PMID:2834197  
A:Accession: S02452  
A:Molecule type: mRNA  
A:Residues: 1-239 <SET>  
R:Cleary, M.L.; Smith, S.D.; Sklar, J.  
Cell 47, 19-28, 1986  
A:Title: Cloning and structural analysis of cDNAs for bcl-2 and a hybrid bcl-2/immunoglobul  
A:Reference number: A24428; MUID:87002488; PMID:2875799  
A:Accession: A24428  
A:Molecule type: mRNA  
A:Residues: 1-58, 'T', '60-116, 'R', '118-239 <CLE>  
A:Cross-references: GB:M14745; NID:g179370; PIDN:AAA35591.1; PID:g179371  
R:Hua, C.; Zorn, S.; Jensen, J.P.; Coupland, R.W.; Ko, H.S.; Wright, J.J.; Bakhshi, A.  
Oncogene Res. 2, 263-275, 1988  
A:Title: Consequences of the t(14;18) chromosomal translocation in follicular lymphoma:  
A:Reference number: A27622; MUID:88217344; PMID:3285301  
A:Accession: A27622  
A:Molecule type: mRNA  
A:Residues: 1-58, 'T', '60-239 <HUA>  
A:Accession: B27622  
A:Molecule type: DNA  
A:Residues: 1-6, 'S', '8-58, 'T', '60-128, 'C', '130-239 <HUA2>  
A:Note: the sequence was determined from the germline gene  
C:Comment: Constitutive expression of BCL2 following t(14;18) chromosomal translocation  
C:Genetics:  
A:Gene: GDB:BCL2  
A:Cross-references: GDB:119031; OMIM:151430  
A:Map position: 18q21.3-18q21.3  
C:Function:  
A:Description: blocks apoptosis in hematopoietic cells  
C:Superfamily: bcl apoptosis regulator, inhibitory type  
C:Keywords: alternative splicing; apoptosis; B-cell lymphoma; follicular lymphoma; proto

Query Match 38.1%; Score 464; DB 1; Length 239;  
Best Local Similarity 44.5%; Pred. No. 4e-33;  
Matches 106; Conservative 30; Mismatches 68; Indels 34; Gaps 8;

Qy 5 NREIWDPLSYKLSQKYSWSQFSDVENRTEAPEETEPEPEETPSAINGNPSW--HLADS 62  
Db 11 NREIWMYIHYKLSQRYEWDA-GDV-----GAAPPGAAP---AFGIFSQFGHTPHPAAS 62  
Qy 63 -----PAVNGATGSHSSLDAREVPMVAQALREAGDEPELRYRRFAFSDLTLS 110  
Db 63 RDPVARTSPLOTPAAPGA-----AAGPALSPPV-PVHLLTLRAGDDFSRRYRDFAEWSS 117  
Qy 111 QLHITPGTAYQSFQVNNELFRDGVNMGRIIVAFPSFGALCVESVDKEMQVLSVRIASWM 170  
Db 118 QLHLTPPTARGFATVVEELFRDGVNMGRIIVAFPEFGVCMCVESVNRMSPLVDNIALWM 177  
Qy 171 ATYLNHLEPWIQNGGWDTFVDLYGNNAEESRKQERFNRWELTGMTVAGVVLGS 228  
Db 178 TEYLNRHLHTWIQNGGWDFAVELYGPSM-----RPLDFSWLSLTKLLSLALVGA 228

RESULT 15  
153744  
Gene bcl-2 protein - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 29-May-1998 #sequence\_revision 29-May-1998 #text\_change 09-Jul-2004  
C:Accession: I53744  
R:Sato, T.; Irie, S.; Krajewski, S.; Reed, J.C.  
Gene 140, 291-292, 1994  
A:Title: Cloning and sequencing of a cDNA encoding the rat Bcl-2 protein.

A;Reference number: I53744; MUID:94193015; PMID:8144041  
 A;Accession: I53744  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: mRNA  
 A;Residues: 1-236 <RES>  
 A;Cross-references: UNIPROT:P49950; GB:L14680; NID:g408946; PIDN:AAA53662.1; PID:g408947  
 C;Genetics:  
 A;Gene: bcl-2  
 C;Superfamily: bcl apoptosis regulator, inhibitory type

Query Match 37.8%; Score 460.5; DB 2; Length 236;  
 Best Local Similarity 43.4%; Pred. NO. 7.9e-33;  
 Matches 102; Conservative 34; Mismatches 68; Indels 31; Gaps 6;  
 Qy 5 NRELVDPLSYKLSQKGYSMQSFSDVENRTEAPEETEPEP--ETPSAINGNPSWHLADS 62  
 Db 11 NREIVMKYIHYKLSQGYEWD-----TGDEDSAPLRAAPTGGIFSQPESN--RT 58  
 Qy 63 PAVNGATGHSSSLDAREVIPMA-----AVKQALREAGDEFEELRYRRAPFSDLTSOLH 113  
 Db 59 PAVHRDTAARTS-PLRPLVANAGPALSPVPVHLLTLRRAGDDPSRRYRRDFAEMSSOLH 117  
 Qy 114 ITPGTAYOSFEQVNNELFRDGVNMGRIIVAFPSFGALCVESVDKEMQVLSRIASWMATY 173  
 Db 118 LTPPTARGRFATVVEELFRDGVNMGRIIVAFPEFGGVMCGSVNREMSPLVDNIALMMTEY 177  
 Qy 174 LNDHLEPMIOENGCHDTFVDLYGNNAAESRKQERFNRFLTGMTVAGVVLGGS 228  
 Db 178 LNRHLHTWIQDNGGWDFAVELYGPSM-----RPLDFSWLSLKTLLSLALVGA 225

Search completed: March 31, 2005, 00:31:34  
 Job time : 32.4692 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2005, 00:25:06 ; Search time 112.019 Seconds  
(without alignments)  
688.690 Million cell updates/sec

Title: US-10-049-822a-2  
Perfect score: 1218  
Sequence: 1 MSQSNRELVDLSYKLSQK.....FLGTMVAGVLLGSLFSRK 233

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1407402 seqs, 331100923 residues

Total number of hits satisfying chosen parameters: 1407402

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
20: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	1196	98.2	233	15	US-10-402-017-4
2	1190	97.7	233	14	US-10-101-482-14
3	1184	97.2	233	9	US-09-734-846-2
4	1184	97.2	233	9	US-09-952-278-6
5	1184	97.2	233	14	US-10-072-830-4
6	1184	97.2	233	14	US-10-169-223-10
7	1184	97.2	233	15	US-10-302-262-2
8	1184	97.2	233	15	US-10-116-275-171
9	1184	97.2	233	16	US-10-450-366-5
10	1184	97.2	233	16	US-10-659-705-8
11	1184	97.2	411	16	US-10-792-517-2
12	1090	89.5	219	15	US-10-402-017-10
13	1079	88.6	219	15	US-10-402-017-12
					Sequence 4, Appli
					Sequence 14, Appl
					Sequence 2, Appli
					Sequence 6, Appli
					Sequence 4, Appli
					Sequence 10, Appl
					Sequence 2, Appli
					Sequence 171, App
					Sequence 5, Appli
					Sequence 8, Appli
					Sequence 2, Appli
					Sequence 12, Appl

14	1075	88.3	212	14	US-10-169-223-14	Sequence 14, Appli
15	1061	87.1	485	16	US-10-792-517-8	Sequence 8, Appli
16	979	80.4	235	14	US-10-208-155-2	Sequence 2, Appli
17	967	79.4	199	15	US-10-402-017-8	Sequence 8, Appli
18	949	77.9	185	9	US-09-864-761-40954	Sequence 40954, A
19	917	75.3	229	16	US-10-659-705-7	Sequence 7, Appli
20	866	71.1	179	15	US-10-402-017-6	Sequence 6, Appli
21	808.5	66.4	170	15	US-10-003-632C-7	Sequence 7, Appli
22	802.5	65.9	170	9	US-09-952-278-8	Sequence 8, Appli
23	751	61.7	152	14	US-10-158-769-2	Sequence 2, Appli
24	732	60.1	190	9	US-09-952-278-2	Sequence 2, Appli
25	588	48.3	237	16	US-10-659-705-5	Sequence 5, Appli
26	515.5	42.3	204	16	US-10-659-705-6	Sequence 6, Appli
27	488	40.1	233	16	US-10-659-705-3	Sequence 3, Appli
28	464	38.1	239	14	US-10-277-693A-10	Sequence 10, Appl
29	464	38.1	239	15	US-10-003-632C-10	Sequence 10, Appl
30	464	38.1	239	15	US-10-003-632C-13	Sequence 13, Appl
31	461	37.8	239	8	US-08-726-211-5	Sequence 5, Appli
32	461	37.8	239	10	US-09-993-420A-8	Sequence 8, Appli
33	461	37.8	239	14	US-10-101-482-12	Sequence 12, Appl
34	461	37.8	239	14	US-10-072-830-2	Sequence 2, Appli
35	461	37.8	239	14	US-10-141-618-12	Sequence 12, Appl
36	461	37.8	239	14	US-10-053-645A-21	Sequence 21, Appl
37	461	37.8	239	15	US-10-387-961A-5	Sequence 5, Appli
38	461	37.8	239	15	US-10-003-632C-1	Sequence 1, Appli
39	461	37.8	239	15	US-10-003-632C-3	Sequence 3, Appli
40	461	37.8	239	15	US-10-148-953A-1	Sequence 1, Appli
41	461	37.8	239	16	US-10-450-366-4	Sequence 4, Appli
42	461	37.8	239	16	US-10-770-668-16	Sequence 16, Appl
43	458.5	37.6	236	13	US-10-087-192-1953	Sequence 1953, Ap
44	458	37.6	239	15	US-10-148-953A-4	Sequence 4, Appli
45	458	37.6	239	15	US-10-148-953A-4	Sequence 4, Appli

## ALIGNMENTS

### RESULT 1

US-10-402-017-4  
; Sequence 4, Application US/10402017  
; Publication No. US20030219871A1  
; GENERAL INFORMATION: Heiko MEENTIS and Martin FUSSENEGGER  
; APPLICANT: Barbara ENEKEL, Host cells having improved survival properties and methods to gen  
; TITLE OF INVENTION: such cells  
; FILE REFERENCE: Case 1/1314  
; CURRENT APPLICATION NUMBER: US/10/402,017  
; CURRENT FILING DATE: 2003-03-28  
; PRIOR APPLICATION NUMBER: US 60/369,307  
; PRIOR APPLICATION NUMBER: April 2, 2002  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 233  
; TYPE: PRT  
; ORGANISM: Cricetus griseus  
US-10-402-017-4

Query Match	98.2%	Score	1196;	DB	15;	Length	233;
Best Local Similarity	98.3%	Pred. No.	6.7e-111;				
Matches	229;	Conservative	2;	Mismatches	2;	Indels	0;
Gaps	0;						
Qy	1	MSQSNRELVDLSYKLSQKYSWSQSDVEENRTEAPEETEPEPTPSAINGNPSWHLA	60				
Db	1	MSQSNRELVDLSYKLSQKYSWSQSDVEENRTEAPEETEPEPTPSAINGNPSWHLA	60				
Qy	61	DSPAVNGATGHSSSLDAREVIPMAAVKQALREAGDEFEELRYRRAFSDLTSLQHTPGTAY	120				
Db	61	DSPAVNGATGHSSSLDAREVIPMAAVKQALREAGDEFEELRYRRAFSDLTSLQHTPGTAY	120				
Qy	121	QSPFQVNNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLVSRIASWMTATYLNHLEP	180				
Db	121	QSPFQVNNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLVSRIASWMTATYLNHLEP	180				

Qy 181 WIQENGWDTFVDLYGNNAAESRKQGRFNRWFLTGMTVAGVLLGSLFSRK 233  
Db 181 WIQENGWDTFVDLYGNNAAESRKQGRFNRWFLTGMTVAGVLLGSLFSRK 233

RESULT 2  
US-10-101-482-14  
; Sequence 14, Application US/10101482  
; Publication No. US2003000837A1  
; GENERAL INFORMATION:  
; APPLICANT: KIEFER, MICHAEL C.  
; BARR, PHILIP J.  
; TITLE OF INVENTION: NOVEL APOPTOSIS-MODULATING PROTEINS, DNA  
; ENCODING THE PROTEINS AND METHODS OF USE THEREOF  
; NUMBER OF SEQUENCES: 22  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORRISON & FOERSTER  
; STREET: 755 Page Mill Road  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94304-1018  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/101,482  
; FILING DATE: 18-Mar-2002  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/320,157  
; FILING DATE: 07-OCT-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: LEHNHARDT, SUSAN K.  
; REGISTRATION NUMBER: 33,943  
; REFERENCE/DOCKET NUMBER: 23647-20007.20  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 813-5600  
; TELEFAX: (415) 494-0792  
; TELEX: 706141  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 233 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; SEQUENCE DESCRIPTION: SEQ ID NO: 14:  
US-10-101-482-14

Query Match 97.7%; Score 1190; DB 14; Length 233;  
Best Local Similarity 97.9%; Pred. No. 2.7e-110;  
Matches 228; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MSQSNRELVDPLSYKLSQKGYNSQFSVDVEENRTEAPEETEPEPTPSAINGNPSWHLA 60  
Db 1 MSQSNRELVDPLSYKLSQKGYNSQFSVDVEENRTEAPEETEPEPTPSAINGNPSWHLA 60

Qy 61 DSPAVNGATGHSSSLDAREVIPMAAVKQALREAGDEFEFLRYRRAFSDLTSLQHIPTPGTAY 120  
Db 61 DSPAVNGATGHSSSLDAREVIPMAAVKQALREAGDEFEFLRYRRAFSDLTSLQHIPTPGTAY 120

Qy 121 QSFQVNVNELFRDGVNMGRIVAFFSFGGALCVESVDKEMQVLVSRIASWMATYLNHLEP 180  
Db 121 QSFQVNVNELFRDGVNMGRIVAFFSFGGALCVESVDKEMQVLVSRIASWMATYLNHLEP 180

Qy 181 WIQENGWDTFVDLYGNNAAESRKQGRFNRWFLTGMTVAGVLLGSLFSRK 233  
Db 181 WIQENGWDTFVDLYGNNAAESRKQGRFNRWFLTGMTVAGVLLGSLFSRK 233

RESULT 3  
US-09-734-846-2  
; Sequence 2, Application US/09734846  
; Patent No. US20010007025A1  
; GENERAL INFORMATION:  
; APPLICANT: Bennett, C. Frank  
; APPLICANT: Dean, Nicholas M.  
; APPLICANT: Monia, Brett P.  
; APPLICANT: Nickoloff, Brian J.  
; APPLICANT: Zhang, Qingqing  
; TITLE OF INVENTION: Antisense Modulation of bcl-x Expression  
; FILE REFERENCE: ISPH-0528  
; CURRENT APPLICATION NUMBER: US/09/734,846  
; CURRENT FILING DATE: 2000-12-12  
; PRIOR APPLICATION NUMBER: 09/277,020  
; PRIOR FILING DATE: 1998-03-26  
; PRIOR APPLICATION NUMBER: 09/167,921  
; PRIOR FILING DATE: 1998-10-07  
; PRIOR APPLICATION NUMBER: 09/323,743  
; PRIOR FILING DATE: 1999-06-02  
; NUMBER OF SEQ ID NOS: 74  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 233  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-734-846-2

Query Match 97.2%; Score 1184; DB 9; Length 233;  
Best Local Similarity 97.4%; Pred. No. 1.1e-109;  
Matches 227; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MSQSNRELVDPLSYKLSQKGYNSQFSVDVEENRTEAPEETEPEPTPSAINGNPSWHLA 60  
Db 1 MSQSNRELVDPLSYKLSQKGYNSQFSVDVEENRTEAPEETEPEPTPSAINGNPSWHLA 60

Qy 61 DSPAVNGATGHSSSLDAREVIPMAAVKQALREAGDEFEFLRYRRAFSDLTSLQHIPTPGTAY 120  
Db 61 DSPAVNGATGHSSSLDAREVIPMAAVKQALREAGDEFEFLRYRRAFSDLTSLQHIPTPGTAY 120

Qy 121 QSFQVNVNELFRDGVNMGRIVAFFSFGGALCVESVDKEMQVLVSRIASWMATYLNHLEP 180  
Db 121 QSFQVNVNELFRDGVNMGRIVAFFSFGGALCVESVDKEMQVLVSRIASWMATYLNHLEP 180

Qy 181 WIQENGWDTFVDLYGNNAAESRKQGRFNRWFLTGMTVAGVLLGSLFSRK 233  
Db 181 WIQENGWDTFVDLYGNNAAESRKQGRFNRWFLTGMTVAGVLLGSLFSRK 233

RESULT 4  
US-09-952-278-6  
; Sequence 6, Application US/09952278  
; Patent No. US20020137182A1  
; GENERAL INFORMATION:  
; APPLICANT: Thompson, Craig B.  
; Boise, Lawrence H.  
; TITLE OF INVENTION: Vertebrate Apoptosis Gene:  
; Compositions and Methods  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Arnold, White & Durkee  
; STREET: 321 No. US20020137182A1th Clark Street, Suite 800  
; CITY: Chicago  
; STATE: IL  
; COUNTRY: USA  
; ZIP: 60610  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/952,278





NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 233  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: BCL2 proteins  
US-10-659-705-8

Query Match 97.2%; Score 1184; DB 16; Length 233;  
Best Local Similarity 97.4%; Pred. No. 1.1e-109;  
Matches 227; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
  
Qy 1 MSQSNRELVDVFLSYKLSQKGYSSQFSDVVENRTEAPEETEPEPTPSAINGNPSWHLA 60  
Db 1 MSQSNRELVDVFLSYKLSQKGYSSQFSDVVENRTEAPEETEPEPTPSAINGNPSWHLA 60  
  
Qy 61 DSPAVNGATGHSSSLDAREVIPMAAVKQALREAGDEFEELRYRRAFSDLTSLQHTTPGTAY 120  
Db 61 DSPAVNGATGHSSSLDAREVIPMAAVKQALREAGDEFEELRYRRAFSDLTSLQHTTPGTAY 120  
  
Qy 121 QSFEQVNNELFRDGVNNGRIVAFPSFGGALCVESVDKEMQVLVSRIASWMTATYLNHLEP 180  
Db 121 QSFEQVNNELFRDGVNNGRIVAFPSFGGALCVESVDKEMQVLVSRIASWMTATYLNHLEP 180  
  
Qy 181 WIQENGWDTTFVDLYGNNAAESRKQGRFNRNFWLTGMTVAGVLLGSLFSRK 233  
Db 181 WIQENGWDTTFVDLYGNNAAESRKQGRFNRNFWLTGMTVAGVLLGSLFSRK 233

RESULT 11  
US-10-792-517-2  
; Sequence 2, Application US/10792517  
; Publication No. US20040152179A1  
; GENERAL INFORMATION:  
; APPLICANT: Youle et al.  
; TITLE OF INVENTION: RECEPTOR-MEDIATED UPTAKE OF AN EXTRACELLULAR BCL-XL  
; FILE REFERENCE: 4239-55417  
; CURRENT APPLICATION NUMBER: US/10/792,517  
; CURRENT FILING DATE: 2004-03-02  
; PRIOR APPLICATION NUMBER: US/09/639,245  
; PRIOR FILING DATE: 2000-08-15  
; PRIOR APPLICATION NUMBER: 60/149,220  
; PRIOR FILING DATE: 1999-08-16  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 411  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: genetic fusion  
US-10-792-517-2

Query Match 97.2%; Score 1184; DB 16; Length 411;  
Best Local Similarity 97.4%; Pred. No. 2.3e-109;  
Matches 227; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
  
Qy 1 MSQSNRELVDVFLSYKLSQKGYSSQFSDVVENRTEAPEETEPEPTPSAINGNPSWHLA 60  
Db 21 MSQSNRELVDVFLSYKLSQKGYSSQFSDVVENRTEAPEETEPEPTPSAINGNPSWHLA 80  
  
Qy 61 DSPAVNGATGHSSSLDAREVIPMAAVKQALREAGDEFEELRYRRAFSDLTSLQHTTPGTAY 120  
Db 81 DSPAVNGATGHSSSLDAREVIPMAAVKQALREAGDEFEELRYRRAFSDLTSLQHTTPGTAY 140  
  
Qy 121 QSFEQVNNELFRDGVNNGRIVAFPSFGGALCVESVDKEMQVLVSRIASWMTATYLNHLEP 180  
Db 141 QSFEQVNNELFRDGVNNGRIVAFPSFGGALCVESVDKEMQVLVSRIASWMTATYLNHLEP 200  
  
Qy 181 WIQENGWDTTFVDLYGNNAAESRKQGRFNRNFWLTGMTVAGVLLGSLFSRK 233

Db 201 WIQENGWDTTFVDLYGNNAAESRKQGRFNRNFWLTGMTVAGVLLGSLFSRK 253  
  
RESULT 12  
US-10-402-017-10  
; Sequence 10, Application US/10402017  
; Publication No. US20030219871A1  
; GENERAL INFORMATION:  
; APPLICANT: Barbara ENENKEL, Heiko MEENTS and Martin FUSENEGGER  
; TITLE OF INVENTION: Host cells having improved survival properties and methods to gen  
; FILE REFERENCE: Case 1/1314  
; CURRENT APPLICATION NUMBER: US/10/402,017  
; CURRENT FILING DATE: 2003-03-28  
; PRIOR APPLICATION NUMBER: US 60/369,307  
; PRIOR APPLICATION NUMBER: April 2, 2002  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 10  
; LENGTH: 219  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Deletion mutant of SEQ ID NO:4 (del166-83)  
US-10-402-017-10

Query Match 89.5%; Score 1090; DB 15; Length 219;  
Best Local Similarity 91.0%; Pred. No. 2.4e-100;  
Matches 212; Conservative 2; Mismatches 5; Indels 14; Gaps 1;  
  
Qy 1 MSQSNRELVDVFLSYKLSQKGYSSQFSDVVENRTEAPEETEPEPTPSAINGNPSWHLA 60  
Db 1 MSQSNRELVDVFLSYKLSQKGYSSQFSDVVENRTEAPEETEPEPTPSAINGNPSWHLA 60  
  
Qy 61 DSPAVNGATGHSSSLDAREVIPMAAVKQALREAGDEFEELRYRRAFSDLTSLQHTTPGTAY 120  
Db 61 DSPAVNGATGHSSSLDAREVIPMAAVKQALREAGDEFEELRYRRAFSDLTSLQHTTPGTAY 106  
  
Qy 121 QSFEQVNNELFRDGVNNGRIVAFPSFGGALCVESVDKEMQVLVSRIASWMTATYLNHLEP 180  
Db 107 QSFEQVNNELFRDGVNNGRIVAFPSFGGALCVESVDKEMQVLVSRIASWMTATYLNHLEP 166  
  
Qy 181 WIQENGWDTTFVDLYGNNAAESRKQGRFNRNFWLTGMTVAGVLLGSLFSRK 233  
Db 167 WIQENGWDTTFVDLYGNNAAESRKQGRFNRNFWLTGMTVAGVLLGSLFSRK 219

RESULT 13  
US-10-402-017-12  
; Sequence 12, Application US/10402017  
; Publication No. US20030219871A1  
; GENERAL INFORMATION:  
; APPLICANT: Barbara ENENKEL, Heiko MEENTS and Martin FUSENEGGER  
; TITLE OF INVENTION: Host cells having improved survival properties and methods to gen  
; FILE REFERENCE: Case 1/1314  
; CURRENT APPLICATION NUMBER: US/10/402,017  
; CURRENT FILING DATE: 2003-03-28  
; PRIOR APPLICATION NUMBER: US 60/369,307  
; PRIOR APPLICATION NUMBER: April 2, 2002  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 12  
; LENGTH: 219  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Deletion mutant of SEQ ID NO:4 (del126-83)  
US-10-402-017-12

Query Match 88.6%; Score 1079; DB 15; Length 219;  
Best Local Similarity 91.0%; Pred. No. 3e-99;

[illegible]

```

RESULT 14
US-10-169-223-14
; Sequence 14, Application US/10169223
; Publication No. US20030152946A1
; GENERAL INFORMATION:
; APPLICANT: SHIMIZU, Shigemi
; APPLICANT: TSUJIMOTO, Yoshihide
; TITLE OF INVENTION: BH4-Fused Polypeptides
; FILE REFERENCE: 1422-0537P
; CURRENT APPLICATION NUMBER: US/10/169,223
; CURRENT FILING DATE: 2002-11-05
; PRIOR APPLICATION NUMBER: JP 11-371449
; PRIOR FILING DATE: 1999-12-27
; PRIOR APPLICATION NUMBER: PCT/JP00/09274
; PRIOR FILING DATE: 2000-12-26
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: CDS of Synthesized DNA for mutant bcl-xL
US-10-169-223-14

```

Query Match	88.3%	Score 1075;	DB 14;	Length 212;
Best Local Similarity	97.2%;	Pred. No. 7.1e-99;		
Matches 205;	Conservative 2;	Mismatches 4;	Indels 0;	Gaps 0;
Qy	23	SWSQFSDVEENRTAPEETEPEERTPSAINGNPSWHLADSPAVNGATGCHSSSLDAREVIP	82	
Db	2	SWSQFSDVEENRTAPECTESEMTPSAINGNPSWHLADSPAVNGATGCHSSSLDAREVIP	61	
Qy	83	MAAVKQALREAGDBFEIURYRAPSDLTSQLHTTPTGATAYQSQEVVYNELFRDGVNNGRIVA	142	
Db	62	MAAVKQALREAGDBFEIURYRAFSDLTSQLHTTPTGATAYQSQEVVYNELFRDGVNNGRIVA	121	
Qy	143	FFSFGGALCVESVDKEMQVLSYRSRTASMMATYLNHDLEPWIQENGWDTFVPLYGNNAAAE	202	
Db	122	FFSFGGALCVESVDKEMQVLSYRTAAMMATYLNHDLEPWIQENGWDTFVPLYGNNAAAE	181	
Qy	203	SRKQGERPNRWPLTGMTVAGVVLGSLFSRK	233	
Db	182	SRKQGERPNRWPLTGMTVAGVVLGSLFSRK	212	

RESULT 15  
US-10-792-517-8  
; Sequence 8, Application US/10792517  
; Publication No. US20040152179A1  
; GENERAL INFORMATION:  
; APPLICANT: Youle et al.  
; TITLE OF INVENTION: RECEPTOR-MEDIATED UPTAKE OF AN EXTRACELLULAR BCL-XL  
; TITLE OF INVENTION: FUSION PROTEIN INHIBITS APOPTOSIS

```

; FILE REFERENCE: 4239-55417
; CURRENT APPLICATION NUMBER: US/10/792,517
; CURRENT FILING DATE: 2004-03-02
; PRIOR APPLICATION NUMBER: US/09/639,245
; PRIOR FILING DATE: 2000-08-15
; PRIOR APPLICATION NUMBER: 60/149,220
; PRIOR FILING DATE: 1999-08-16
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 485
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: genetic fusion
US-10-792-517-8

```

Query Match	87.1%	Score 1061	DB 16	Length 485
Best Local Similarity	97.1%	Pred. No. 5.6e-97		
Matches 203	Conservative	2	Mismatches 4	Indels 0
Gaps				
Qy	1	MSQSNRELVDFLSYKLSQKGYSWSPDSFDEENRTEAPEETEPRETPSAINGNPSWHLA	60	
Db	277	MSQSNRELVDFLSYKLSQKGYSWSPDSFDEENRTEAPEGTESEMETPSAINGNPSWHLA	336	
Qy	61	DSPAVNGATCHSSLDAREVIPMAAVKQALREAGDEFELRYRRAFSDLTSQLHTTPTGAY	120	
Db	337	DSPAVNGATAHSSLDAREVIPMAAVKQALREAGDEFELRYRRAFSDLTSQLHTTPTGAY	396	
Qy	121	QSFEQVNNELPRGVNNGRIVAFPSFGGALCVESVDKEMQVLSRIASWWTATYLNHDHLEP	180	
Db	397	QSFEQVNNELPRGVNNGRIVAFPSFGGALCVESVDKEMQVLSRIAAWWTATYLNHDHLEP	456	
Qy	181	WIENGWDFTFDLYGNNAEAERKGOER	209	
Db	457	WIENGWDFTFVELYGNNAEAERKGOER	485	

Search completed: March 31, 2005, 00:34:54  
Job time : 115.019 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2005, 00:23:16 ; Search time 46.6 Seconds  
(without alignments)  
373.245 Million cell updates/sec

Title: US-10-049-822a-2  
Perfect score: 1218  
Sequence: 1 MSQSNRELVDLSYKLSQK.....FLTGMTVAGVLLGSLFSRK 233

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A COMB.pep:\*  
2: /cgn2\_6/ptodata/1/iaa/5B COMB.pep:\*  
3: /cgn2\_6/ptodata/1/iaa/6A COMB.pep:\*  
4: /cgn2\_6/ptodata/1/iaa/6B COMB.pep:\*  
5: /cgn2\_6/ptodata/1/iaa/PTUS COMB.pep:\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1190	97.7	233	1	US-08-333-565-59
2	1190	97.7	233	1	US-08-471-058-14
3	1190	97.7	233	2	US-08-661-479-59
4	1190	97.7	233	3	US-08-471-057-14
5	1190	97.7	233	3	US-08-481-739-2
6	1190	97.7	233	4	US-08-470-865-14
7	1190	97.7	233	4	US-08-155-327G-12
8	1190	97.7	233	5	PCT-US94-07089-7
9	1184	97.2	233	1	US-08-081-448-6
10	1184	97.2	233	1	US-08-607-269-24
11	1184	97.2	233	2	US-08-470-670A-7
12	1184	97.2	233	3	US-08-167-921-2
13	1184	97.2	233	3	US-08-323-743-2
14	1184	97.2	233	3	US-08-461-511A-7
15	1184	97.2	233	5	PCT-US95-04600-24
16	1184	97.2	411	4	US-09-639-245-2
17	1177	96.6	233	3	US-09-271-014A-6
18	1128	92.6	225	3	US-09-101-519-1
19	1061	87.1	485	4	US-09-639-245-8
20	979	80.4	235	4	US-08-899-367-2
21	808.5	66.4	170	5	PCT-US94-07089-9
22	808.5	66.4	186	4	US-09-949-016-10416
23	802.5	65.9	170	1	US-08-081-448-8
24	802.5	65.9	170	2	US-08-470-670A-9
25	802.5	65.9	170	3	US-08-461-511A-9
26	802.5	65.9	170	3	US-09-271-014A-8
27	732	60.1	190	1	US-08-081-448-2

28	732	60.1	190	2	US-08-470-670A-2	Sequence 2, Appli
29	732	60.1	190	3	US-08-461-511A-2	Sequence 2, Appli
30	732	60.1	190	3	US-09-271-014A-2	Sequence 2, Appli
31	732	60.1	190	5	PCT-US94-07089-2	Sequence 2, Appli
32	599	49.2	121	2	US-08-470-670A-15	Sequence 15, Appl
33	599	49.2	121	3	US-08-461-511A-15	Sequence 15, Appl
34	510	41.9	109	2	US-08-470-670A-11	Sequence 11, Appl
35	510	41.9	109	3	US-08-461-511A-11	Sequence 11, Appl
36	497	40.8	233	1	US-08-607-269-23	Sequence 23, Appl
37	497	40.8	233	5	PCT-US95-04600-23	Sequence 23, Appl
38	464	38.1	239	1	US-08-112-208C-10	Sequence 10, Appl
39	464	38.1	239	1	US-08-248-819A-10	Sequence 10, Appl
40	464	38.1	239	2	US-08-337-646A-10	Sequence 10, Appl
41	464	38.1	239	2	US-08-856-531-10	Sequence 10, Appl
42	464	38.1	239	2	US-08-856-034-10	Sequence 10, Appl
43	464	38.1	239	3	US-09-127-048-8	Sequence 8, Appli
44	464	38.1	239	3	US-08-927-326-10	Sequence 10, Appl
45	464	38.1	239	4	US-09-379-820A-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1  
US-08-333-565-59  
; Sequence 59, Application US/08333565  
; Patent No. 5622852  
; GENERAL INFORMATION:  
; APPLICANT: KORSMEYER, Stanley J.  
; TITLE OF INVENTION: Rcl-x/Bcl-2 ASSOCIATED CELL DEATH  
; TITLE OF INVENTION: REGULATOR  
; NUMBER OF SEQUENCES: 59  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Kourie and Crew  
; STREET: 379 Lytton Avenue  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: US  
; ZIP: 94301  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/333,565  
; FILING DATE: 31-OCT-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith, William M  
; REGISTRATION NUMBER: 30,223  
; REFERENCE/DOCKET NUMBER: 15726A-000700  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 326-2400  
; TELEFAX: (415) 326-2422  
; INFORMATION FOR SEQ ID NO: 59:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 233 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
US-08-333-565-59

Query Match 97.7%; Score 1190; DB 1; Length 233;  
Best Local Similarity 97.9%; Pred. No. 5.5e-125;  
Matches 228; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
Qy 1 MSQSNRELVDLSYKLSQKSWQSFDVEENRTEAPEETEPEETPSAINGNPSWHLA 60  
Db 1 MSQSNRELVDLSYKLSQKSWQSFDVEENRTEAPEETEPEETPSAINGNPSWHLA 60  
Qy 61 DSPAVNGATGHSSSLDAREVIPMAAVKQALREAGDEFELRYRRAFSDLTSLQHTPTGAY 120

Db 61 DSPAVNGATGHSSSLDAREVIPMAAVKQALREAGDEFELRYRRAFSDLTSQLHTPGTAY 120  
Qy 121 QSPFQVVNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLSRIASWMTATYLNHLEP 180  
Db 121 QSPFQVVNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLSRIASWMTATYLNHLEP 180  
Qy 181 WIQENGWDTFVDLYGNNAAESKGOERFNRWFLTGMTVAGVVLLGSLFSRK 233  
Db 181 WIQENGWDTFVELYGNNAAESKGOERFNRWFLTGMTVAGVVLLGSLFSRK 233  
RESULT 2  
US-08-471-058-14  
; Sequence 14, Application US/08471058  
; Patent No. 5770443  
; GENERAL INFORMATION:  
; APPLICANT: Kiefer, Michael C.  
; APPLICANT: Kiefer, Philip J.  
; TITLE OF INVENTION: NOVEL APOPTOSIS MODULATING  
; TITLE OF INVENTION: PROTEINS, DNA ENCODING THE PROTEINS AND METHODS OF USE  
; TITLE OF INVENTION: THEREOF  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORRISON & FOERSTER  
; STREET: 755 PAGE MILL ROAD  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304-1018  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Dikette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/471,058  
; FILING DATE: 06-JUN-1995  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/320,157  
; FILING DATE: 07-OCT-1994  
; APPLICATION NUMBER: 08/160,067  
; FILING DATE: 30-NOV-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lehnhardt, Susan K  
; REGISTRATION NUMBER: 33,943  
; REFERENCE/DOCKET NUMBER: 23647-20007.12  
; TELEPHONE: 415-813-5600  
; TELEFAX: 415-494-0792  
; TELEX: 706141  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 233 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-471-058-14  
Query Match 97.7%; Score 1190; DB 1; Length 233;  
Best Local Similarity 97.9%; Pred. No. 5.5e-125;  
Matches 228; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
Qy 1 MSQSNRELVDVFLSYKLSQKGYNSQSFSDVEENRTEAPEETEPEPTPSAINGNPSWHLA 60  
Db 1 MSQSNRELVDVFLSYKLSQKGYNSQSFSDVEENRTEAPEETEPEPTPSAINGNPSWHLA 60  
Qy 61 DSPAVNGATGHSSSLDAREVIPMAAVKQALREAGDEFELRYRRAFSDLTSQLHTPGTAY 120  
Db 61 DSPAVNGATGHSSSLDAREVIPMAAVKQALREAGDEFELRYRRAFSDLTSQLHTPGTAY 120  
Qy 121 QSPFQVVNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLSRIASWMTATYLNHLEP 180  
Db 121 QSPFQVVNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLSRIASWMTATYLNHLEP 180  
Qy 181 WIQENGWDTFVDLYGNNAAESKGOERFNRWFLTGMTVAGVVLLGSLFSRK 233  
Db 181 WIQENGWDTFVELYGNNAAESKGOERFNRWFLTGMTVAGVVLLGSLFSRK 233

Db 121 QSPFQVVNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLSRIASWMTATYLNHLEP 180  
Qy 181 WIQENGWDTFVDLYGNNAAESKGOERFNRWFLTGMTVAGVVLLGSLFSRK 233  
Db 181 WIQENGWDTFVELYGNNAAESKGOERFNRWFLTGMTVAGVVLLGSLFSRK 233  
RESULT 3  
US-08-661-479-59  
; Sequence 59, Application US/08661479  
; Patent No. 5834209  
; GENERAL INFORMATION:  
; APPLICANT: KORSMEYER, Stanley J.  
; TITLE OF INVENTION: Bcl-x/Bcl-2 ASSOCIATED CELL DEATH  
; TITLE OF INVENTION: REGULATOR  
; NUMBER OF SEQUENCES: 59  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Kourie and Crew  
; STREET: 379 Lytton Avenue  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: US  
; ZIP: 94301  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/661,479  
; FILING DATE: 11-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/333,565  
; FILING DATE: 31-OCT-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith, William M  
; REGISTRATION NUMBER: 30,223  
; REFERENCE/DOCKET NUMBER: 15726A-000700  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 326-2400  
; TELEFAX: (415) 326-2422  
; INFORMATION FOR SEQ ID NO: 59:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 233 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
US-08-661-479-59  
Query Match 97.7%; Score 1190; DB 2; Length 233;  
Best Local Similarity 97.9%; Pred. No. 5.5e-125;  
Matches 228; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
Qy 1 MSQSNRELVDVFLSYKLSQKGYNSQSFSDVEENRTEAPEETEPEPTPSAINGNPSWHLA 60  
Db 1 MSQSNRELVDVFLSYKLSQKGYNSQSFSDVEENRTEAPEETEPEPTPSAINGNPSWHLA 60  
Qy 61 DSPAVNGATGHSSSLDAREVIPMAAVKQALREAGDEFELRYRRAFSDLTSQLHTPGTAY 120  
Db 61 DSPAVNGATGHSSSLDAREVIPMAAVKQALREAGDEFELRYRRAFSDLTSQLHTPGTAY 120  
Qy 121 QSPFQVVNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLSRIASWMTATYLNHLEP 180  
Db 121 QSPFQVVNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLSRIASWMTATYLNHLEP 180  
Qy 181 WIQENGWDTFVDLYGNNAAESKGOERFNRWFLTGMTVAGVVLLGSLFSRK 233  
Db 181 WIQENGWDTFVELYGNNAAESKGOERFNRWFLTGMTVAGVVLLGSLFSRK 233



```

RESULT 4
US-08-471-057-14
; Sequence 14, Application US/08471057
; Patent No. 6015687
; GENERAL INFORMATION:
; APPLICANT: KIEFER, MICHAEL C.
; APPLICANT: BARR, PHILIP J.
; TITLE OF INVENTION: NOVEL APOPTOSIS-MODULATING PROTEINS, DNA
; TITLE OF INVENTION: ENCODING THE PROTEINS AND METHODS OF USE THEREOF
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/471.057
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/320,157
; FILING DATE: 07-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: LEHNHARDT, SUSAN K.
; REGISTRATION NUMBER: 33,943
; REFERENCE/DOCKET NUMBER: 23647-20007.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 233 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-471-057-14

Query Match 97.7%; Score 1190; DB 3; Length 233;
Best Local Similarity 97.9%; Pred. No. 5.5e-125;
Matches 228; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MSQSNRELVDVFLSKYKLSQKGYSSQFSDVEENRTEAPEETEPEETPSAINGNPSWHLA 60
Db 1 MSQSNRELVDVFLSKYKLSQKGYSSQFSDVEENRTEAPEETEPEETPSAINGNPSWHLA 60
Qy 61 DSPAVNGATGHSSSLDAREVPMVAVKQALREAGDEFELRYRRAFSDLTSLQHIPTGTAY 120
Db 61 DSPAVNGATGHSSSLDAREVPMVAVKQALREAGDEFELRYRRAFSDLTSLQHIPTGTAY 120
Qy 121 QSPFQVNNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLVSRIASWMTATYLNHLEP 180
Db 121 QSPFQVNNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLVSRIASWMTATYLNHLEP 180
Qy 181 WIOENGWDTFVLYGNNAAESRKQGRFNRWFLTGMTVAGVVLGSLFSRK 233
Db 181 WIOENGWDTFVLYGNNAAESRKQGRFNRWFLTGMTVAGVVLGSLFSRK 233

RESULT 5
US-08-481-739-2
; Sequence 2, Application US/08481739
; Patent No. 6143291
; GENERAL INFORMATION:
; APPLICANT: June, Carl H. and Thompson, Craig B.
; TITLE OF INVENTION: METHODS FOR ENHANCING T CELL SURVIVAL

```

```

; TITLE OF INVENTION: BY AUGMENTING BCL-XL PROTEIN LEVELS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/481,739
; FILING DATE: 07-JUNE-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/435,518
; FILING DATE: 04-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: DeConti, Giulio A. (GAD)
; REGISTRATION NUMBER: 31,503
; REFERENCE/DOCKET NUMBER: RPI-034CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 742-4214
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 233 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-481-739-2

Query Match 97.7%; Score 1190; DB 3; Length 233;
Best Local Similarity 97.9%; Pred. No. 5.5e-125;
Matches 228; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MSQSNRELVDVFLSKYKLSQKGYSSQFSDVEENRTEAPEETEPEETPSAINGNPSWHLA 60
Db 1 MSQSNRELVDVFLSKYKLSQKGYSSQFSDVEENRTEAPEETEPEETPSAINGNPSWHLA 60
Qy 61 DSPAVNGATGHSSSLDAREVPMVAVKQALREAGDEFELRYRRAFSDLTSLQHIPTGTAY 120
Db 61 DSPAVNGATGHSSSLDAREVPMVAVKQALREAGDEFELRYRRAFSDLTSLQHIPTGTAY 120
Qy 121 QSPFQVNNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLVSRIASWMTATYLNHLEP 180
Db 121 QSPFQVNNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLVSRIASWMTATYLNHLEP 180
Qy 181 WIOENGWDTFVLYGNNAAESRKQGRFNRWFLTGMTVAGVVLGSLFSRK 233
Db 181 WIOENGWDTFVLYGNNAAESRKQGRFNRWFLTGMTVAGVVLGSLFSRK 233

RESULT 6
US-08-470-865-14
; Sequence 14, Application US/08470865
; Patent No. 6586395
; GENERAL INFORMATION:
; APPLICANT: KIEFER, MICHAEL C.
; APPLICANT: BARR, PHILIP J.
; TITLE OF INVENTION: NOVEL APOPTOSIS-MODULATING PROTEINS, DNA
; TITLE OF INVENTION: ENCODING THE PROTEINS AND METHODS OF USE THEREOF
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1018

```

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470,865
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/320,157
; FILING DATE: 07-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: LEHNHARDT, SUSAN K.
; REGISTRATION NUMBER: 33,943
; REFERENCE/DOCKET NUMBER: 23647-20007.20
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 233 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-470-865-14

Query Match          97.7%; Score 1190; DB 4; Length 233;
Best Local Similarity 97.9%; Pred. No. 5.5e-125;
Matches 228; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MSQSNRELVDFLSYKLSQKGSWSQFSDVEENRTEAPEETEPEETPSAINGNPSWHLA 60
Db 1 MSQSNRELVDFLSYKLSQKGSWSQFSDVEENRTEAPEETEPEETPSAINGNPSWHLA 60

Qy 61 DSPAVNGATGHSSSLDAREVPMAAVKQALREAGDEFELRYRRAFSDLTSQLHTPGTAY 120
Db 61 DSPAVNGATGHSSSLDAREVPMAAVKQALREAGDEFELRYRRAFSDLTSQLHTPGTAY 120

Qy 121 QSPFQVYNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLVSRIASWMATYLNHLEP 180
Db 121 QSPFQVYNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLVSRIASWMATYLNHLEP 180

Qy 181 WIQENGWDITFDVLYGNNAAESRKQGRFNRWFLTGMTVAGVVLGSLFSRK 233
Db 181 WIQENGWDITFDVLYGNNAAESRKQGRFNRWFLTGMTVAGVVLGSLFSRK 233

RESULT 8
PCT-US94-07089-7
; Sequence 7, Application PC/TUS9407089
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Vertebrate Apoptosis Gene:
; TITLE OF INVENTION: Compositions and Methods
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: TX
; COUNTRY: United States of America
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/07089
; FILING DATE: CONCURRENTLY FILED
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/081.448
; FILING DATE: 22 JUNE 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: PARKER, David L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: ARCD090
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512-320-7200
; TELEFAX: 713-789-2679
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 233 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US94-07089-7

Query Match          97.7%; Score 1190; DB 5; Length 233;
Best Local Similarity 97.9%; Pred. No. 5.5e-125;
Matches 228; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MSQSNRELVDFLSYKLSQKGSWSQFSDVEENRTEAPEETEPEETPSAINGNPSWHLA 60
Db 1 MSQSNRELVDFLSYKLSQKGSWSQFSDVEENRTEAPEETEPEETPSAINGNPSWHLA 60

Qy 61 DSPAVNGATGHSSSLDAREVPMAAVKQALREAGDEFELRYRRAFSDLTSQLHTPGTAY 120
Db 61 DSPAVNGATGHSSSLDAREVPMAAVKQALREAGDEFELRYRRAFSDLTSQLHTPGTAY 120

Qy 121 QSPFQVYNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLVSRIASWMATYLNHLEP 180
Db 121 QSPFQVYNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLVSRIASWMATYLNHLEP 180

```

Db 121 QSFQVNVNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLSRIAAMWATYLNHLEP 180  
Qy 181 WIQENGWDTFVDLYGNNAAESKQGRFNRWFLTGMTVAGVLLGSLFSRK 233  
Db 181 WIQENGWDTFVLYGNNAAESKQGRFNRWFLTGMTVAGVLLGSLFSRK 233

## RESULT 9

US-08-081-448-6  
; Sequence 6, Application US/08081448  
; Patent No. 5646008  
; GENERAL INFORMATION:  
; APPLICANT: Thompson, Craig B.  
; APPLICANT: Boise, Lawrence H.  
; TITLE OF INVENTION: Vertebrate Apoptosis Gene:  
; TITLE OF INVENTION: Compositions and Methods  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Arnold, White & Durkee  
; STREET: 321 No. 5646008th Clark Street, Suite 800  
; CITY: Chicago  
; STATE: IL  
; COUNTRY: USA  
; ZIP: 60610

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/081,448  
; FILING DATE: 19930622  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 5646008thrup, Thomas E.  
; REGISTRATION NUMBER: 33,268  
; REFERENCE/DOCKET NUMBER: ARCD090  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-744-0090  
; TELEFAX: 312-755-4489  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 233 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-081-448-6

Query Match 97.2%; Score 1184; DB 1; Length 233;  
Best Local Similarity 97.4%; Pred. No. 2.6e-124;  
Matches 227; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MSQSNRELVDFLSYKLSQKGYSSQFSDVEENTEAPEETEPEETPSAINGNPSWHLA 60  
Db 1 MSQSNRELVDFLSYKLSQKGYSSQFSDVEENTEAPEETEPEETPSAINGNPSWHLA 60  
Qy 61 DSPAVNGATGHSSSLDAREVIPMAVKQALREAGDEPELRYRRAFSDLTSLHITPGTAY 120  
Db 61 DSPAVNGATGHSSSLDAREVIPMAVKQALREAGDEPELRYRRAFSDLTSLHITPGTAY 120  
Qy 121 QSFQVNVNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLSRIAAMWATYLNHLEP 180  
Db 121 QSFQVNVNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLSRIAAMWATYLNHLEP 180  
Qy 181 WIQENGWDTFVDLYGNNAAESKQGRFNRWFLTGMTVAGVLLGSLFSRK 233  
Db 181 WIQENGWDTFVLYGNNAAESKQGRFNRWFLTGMTVAGVLLGSLFSRK 233

## RESULT 10

US-08-607-269-24  
; Sequence 24, Application US/08607269  
; Patent No. 5702897

; GENERAL INFORMATION:  
; APPLICANT: Reed, John C.  
; APPLICANT: Sato, Takaaki  
; TITLE OF INVENTION: Interaction of Proteins Involved in a  
; TITLE OF INVENTION: Cell Death Pathway  
; NUMBER OF SEQUENCES: 29  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Campbell and Flores  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA: US/08/607,269  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA: US 08/226,876  
; APPLICATION NUMBER: US 08/226,876  
; FILING DATE: 13-APR-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Campbell, Cathryn A.  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: P-LJ 9882  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 535-9001  
; TELEFAX: (619) 535-8949  
; INFORMATION FOR SEQ ID NO: 24:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 233 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-08-607-269-24

Query Match 97.2%; Score 1184; DB 1; Length 233;  
Best Local Similarity 97.4%; Pred. No. 2.6e-124;  
Matches 227; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MSQSNRELVDFLSYKLSQKGYSSQFSDVEENTEAPEETEPEETPSAINGNPSWHLA 60  
Db 1 MSQSNRELVDFLSYKLSQKGYSSQFSDVEENTEAPEETEPEETPSAINGNPSWHLA 60  
Qy 61 DSPAVNGATGHSSSLDAREVIPMAVKQALREAGDEPELRYRRAFSDLTSLHITPGTAY 120  
Db 61 DSPAVNGATGHSSSLDAREVIPMAVKQALREAGDEPELRYRRAFSDLTSLHITPGTAY 120  
Qy 121 QSFQVNVNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLSRIAAMWATYLNHLEP 180  
Db 121 QSFQVNVNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLSRIAAMWATYLNHLEP 180  
Qy 181 WIQENGWDTFVDLYGNNAAESKQGRFNRWFLTGMTVAGVLLGSLFSRK 233  
Db 181 WIQENGWDTFVLYGNNAAESKQGRFNRWFLTGMTVAGVLLGSLFSRK 233

## RESULT 11

US-08-470-670A-7  
; Sequence 7, Application US/08470670A  
; Patent No. 5834309  
; Patent No. 5834309 5710045  
; GENERAL INFORMATION:  
; APPLICANT: Thompson, Craig B. B.  
; APPLICANT: Boise, Lawrence H.  
; TITLE OF INVENTION: VERTEBRATE APOPTOSIS GENE:  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Arnold, White & Durkee

STREET: P.O. Box 4433  
 CITY: Houston  
 STATE: Texas  
 COUNTRY: United States of America  
 ZIP: 77210  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/470,670A  
 FILING DATE:  
 CLASSIFICATION: 514  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/081,448  
 FILING DATE: 22-JUN-1993  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Highlander, Steven L.  
 REGISTRATION NUMBER: 37,642  
 REFERENCE/DOCKET NUMBER: ARCD:090--1  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (512) 418-3000  
 TELEFAX: (512) 474-7577  
 INFORMATION FOR SEQ ID NO: 7:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 233 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-470-670A-7

Query Match 97.2%; Score 1184; DB 2; Length 233;  
 Best Local Similarity 97.4%; Pred. No. 2.6e-124;  
 Matches 227; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 MSQSNRELVDVFLSYKLSQKGYWSQFSDVEENRTEAPEETEPEPERETPSAINGNPSWHLA 60  
 DB 1 MSQSNRELVDVFLSYKLSQKGYWSQFSDVEENRTEAPEETEPEPERETPSAINGNPSWHLA 60  
 QY 61 DSPAVNGATGHSSSLDAREVIPMAAVKQALREAGDEFELRYRRAPSDLTSQLHTTPGTAY 120  
 DB 61 DSPAVNGATGHSSSLDAREVIPMAAVKQALREAGDEFELRYRRAPSDLTSQLHTTPGTAY 120  
 QY 121 QSFQVNVNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLVSRIASWMTATYLNHLEP 180  
 DB 121 QSFQVNVNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLVSRIASWMTATYLNHLEP 180  
 QY 181 WIQENGWDTFVDLYGNNAEAESRKQGRFNRWFLTGMTVAGVLLGSLFSRK 233  
 DB 181 WIQENGWDTFVDLYGNNAEAESRKQGRFNRWFLTGMTVAGVLLGSLFSRK 233

RESULT 12  
 US-09-167-921-2  
 ; Sequence 2, Application US/09167921A  
 ; Patent No. 6172216  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bennett, C. Frank  
 ; APPLICANT: Dean, Nicholas M.  
 ; APPLICANT: Monia, Brett P.  
 ; APPLICANT: Nickoloff, Brian J.  
 ; APPLICANT: Zhang, QingQing  
 ; TITLE OF INVENTION: Antisense Modulation of bcl-x Expression  
 ; FILE REFERENCE: ISPH-0324  
 ; CURRENT APPLICATION NUMBER: US/09/167,921A  
 ; CURRENT FILING DATE: 1998-10-07  
 ; NUMBER OF SEQ ID NOS: 50  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 2  
 ; LENGTH: 233  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens

US-09-167-921-2  
 Query Match 97.2%; Score 1184; DB 3; Length 233;  
 Best Local Similarity 97.4%; Pred. No. 2.6e-124;  
 Matches 227; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 MSQSNRELVDVFLSYKLSQKGYWSQFSDVEENRTEAPEETEPEPERETPSAINGNPSWHLA 60  
 DB 1 MSQSNRELVDVFLSYKLSQKGYWSQFSDVEENRTEAPEETEPEPERETPSAINGNPSWHLA 60  
 QY 61 DSPAVNGATGHSSSLDAREVIPMAAVKQALREAGDEFELRYRRAPSDLTSQLHTTPGTAY 120  
 DB 61 DSPAVNGATGHSSSLDAREVIPMAAVKQALREAGDEFELRYRRAPSDLTSQLHTTPGTAY 120  
 QY 121 QSFQVNVNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLVSRIASWMTATYLNHLEP 180  
 DB 121 QSFQVNVNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLVSRIASWMTATYLNHLEP 180  
 QY 181 WIQENGWDTFVDLYGNNAEAESRKQGRFNRWFLTGMTVAGVLLGSLFSRK 233  
 DB 181 WIQENGWDTFVDLYGNNAEAESRKQGRFNRWFLTGMTVAGVLLGSLFSRK 233

RESULT 13  
 US-09-323-743-2  
 ; Sequence 2, Application US/09323743  
 ; Patent No. 6214986  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bennett, C. Frank  
 ; APPLICANT: Dean, Nicholas M.  
 ; APPLICANT: Monia, Brett P.  
 ; APPLICANT: Nickoloff, Brian J.  
 ; APPLICANT: Zhang, QingQing  
 ; TITLE OF INVENTION: Antisense Modulation of bcl-x Expression  
 ; FILE REFERENCE: ISPH-0368  
 ; CURRENT APPLICATION NUMBER: US/09/323,743  
 ; CURRENT FILING DATE: 1999-06-01  
 ; EARLIER APPLICATION NUMBER: 09/277,020  
 ; EARLIER FILING DATE: 1998-03-26  
 ; EARLIER APPLICATION NUMBER: 09/167,921  
 ; EARLIER FILING DATE: 1998-10-07  
 ; NUMBER OF SEQ ID NOS: 66  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 2  
 ; LENGTH: 233  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-323-743-2

Query Match 97.2%; Score 1184; DB 3; Length 233;  
 Best Local Similarity 97.4%; Pred. No. 2.6e-124;  
 Matches 227; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 MSQSNRELVDVFLSYKLSQKGYWSQFSDVEENRTEAPEETEPEPERETPSAINGNPSWHLA 60  
 DB 1 MSQSNRELVDVFLSYKLSQKGYWSQFSDVEENRTEAPEETEPEPERETPSAINGNPSWHLA 60  
 QY 61 DSPAVNGATGHSSSLDAREVIPMAAVKQALREAGDEFELRYRRAPSDLTSQLHTTPGTAY 120  
 DB 61 DSPAVNGATGHSSSLDAREVIPMAAVKQALREAGDEFELRYRRAPSDLTSQLHTTPGTAY 120  
 QY 121 QSFQVNVNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLVSRIASWMTATYLNHLEP 180  
 DB 121 QSFQVNVNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLVSRIASWMTATYLNHLEP 180  
 QY 181 WIQENGWDTFVDLYGNNAEAESRKQGRFNRWFLTGMTVAGVLLGSLFSRK 233  
 DB 181 WIQENGWDTFVDLYGNNAEAESRKQGRFNRWFLTGMTVAGVLLGSLFSRK 233

RESULT 14  
 US-08-461-511A-7  
 ; Sequence 7, Application US/08461511A

Patent No. 6303331  
GENERAL INFORMATION:  
APPLICANT: Thompson, Craig B.B.  
Boise, Lawrence H.  
TITLE OF INVENTION: VERTEBRATE APOPTOSIS GENE: COMPOSITIONS  
AND METHODS  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: United States of America  
ZIP: 77210

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/461,511A  
FILING DATE: 05-Jun-1995  
CLASSIFICATION: UNKNOWN  
ATTORNEY/AGENT INFORMATION:  
NAME: Highlander, Steven L.  
REGISTRATION NUMBER: 37,642  
REFERENCE/DOCKET NUMBER: ARCD:179  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (512) 418-3000  
TELEFAX: (512) 474-7577

INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 233 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 7:  
US-08-461-511A-7

Query Match 97.2%; Score 1184; DB 3; Length 233;  
Best Local Similarity 97.4%; Pred. No. 2.6e-124;  
Matches 227; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
Qy 1 MSQSNRELVDVFLSYKLSQKGYSSQFSQSDVEENRTEAPEETPERETPSAINGNPSWHLA 60  
Db 1 MSQSNRELVDVFLSYKLSQKGYSSQFSQSDVEENRTEAPEETPERETPSAINGNPSWHLA 60  
Qy 61 DSPAVNGATGHSSSLDAREVIPMAAVKQALREAGDEFEFLRYRRAFSDLTSLQHTTPGTAY 120  
Db 61 DSPAVNGATGHSSSLDAREVIPMAAVKQALREAGDEFEFLRYRRAFSDLTSLQHTTPGTAY 120  
Qy 121 QSFQVNVNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLSRIASWMTYLNHLEP 180  
Db 121 QSFQVNVNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLSRIASWMTYLNHLEP 180  
Qy 181 WIQENGWDTFVDLYGNNAAESRKQERFNRWFLTGMTVAGVVLGSLFSRK 233  
Db 181 WIQENGWDTFVDLYGNNAAESRKQERFNRWFLTGMTVAGVVLGSLFSRK 233

RESULT 15  
PCT-US95-04600-24  
Sequence 24, Application PC/TUS9504600  
GENERAL INFORMATION:  
APPLICANT: LA JOLLA CANCER RESEARCH FOUNDATION  
TITLE OF INVENTION: Interaction of Proteins Involved in  
a Cell Death Pathway  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Campbell and Flores  
STREET: 4370 La Jolla Village Drive, Suite 700  
CITY: San Diego  
STATE: California

COUNTRY: USA  
ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/04600  
FILING DATE: 12-APR-1995  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Imbra, Richard J.  
REGISTRATION NUMBER: 37,643  
REFERENCE/DOCKET NUMBER: FP-LJ 1361  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949

INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 233 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
PCT-US95-04600-24

Query Match 97.2%; Score 1184; DB 5; Length 233;  
Best Local Similarity 97.4%; Pred. No. 2.6e-124;  
Matches 227; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
Qy 1 MSQSNRELVDVFLSYKLSQKGYSSQFSQSDVEENRTEAPEETPERETPSAINGNPSWHLA 60  
Db 1 MSQSNRELVDVFLSYKLSQKGYSSQFSQSDVEENRTEAPEETPERETPSAINGNPSWHLA 60  
Qy 61 DSPAVNGATGHSSSLDAREVIPMAAVKQALREAGDEFEFLRYRRAFSDLTSLQHTTPGTAY 120  
Db 61 DSPAVNGATGHSSSLDAREVIPMAAVKQALREAGDEFEFLRYRRAFSDLTSLQHTTPGTAY 120  
Qy 121 QSFQVNVNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLSRIASWMTYLNHLEP 180  
Db 121 QSFQVNVNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLSRIASWMTYLNHLEP 180  
Qy 181 WIQENGWDTFVDLYGNNAAESRKQERFNRWFLTGMTVAGVVLGSLFSRK 233  
Db 181 WIQENGWDTFVDLYGNNAAESRKQERFNRWFLTGMTVAGVVLGSLFSRK 233

Search completed: March 31, 2005, 00:32:37  
Job time : 48.6 secs

Handwritten text, possibly a signature or a list of names, oriented vertically.

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2005, 00:17:06 ; Search time 163.1 Seconds  
(without alignments)  
552.515 Million cell updates/sec

Title: US-10-049-822a-2  
Perfect score: 1218  
Sequence: 1 MSQSNRELVDVFLSYKLSQK.....FLGTMTVAGVLLGSLFSRK 233

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*  
1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	1218	100.0	233	4	AAB73303 Rat wild-
2	1218	100.0	233	7	AD622921 Rat Prote
3	1218	100.0	233	7	AD62491 Rat Prote
4	1210	99.3	233	8	ADQ80678 Mouse ant
5	1206	99.0	233	4	AAB73304 Mutant ra
6	1197	98.3	237	5	ABG78480 Wild type
7	1196	98.2	233	8	ADH52633 Chinese h
8	1190	97.7	233	2	AA68887 Human thy
9	1190	97.7	233	2	AAW05821 Bcl-XL pr
10	1190	97.7	233	2	AAW31530 Human ant
11	1190	97.7	233	7	AD62493 Human Pro
12	1190	97.7	233	8	ADN04261 Antipsori
13	1190	97.7	233	8	ADM45995 Human apo
14	1190	97.7	233	8	ADO19867 Human PRO
15	1190	97.7	233	8	ABM82217 Tumour-as
16	1190	97.7	233	8	ADP54629 Human PRO
17	1190	97.7	233	8	ADP54992 Human PRO
18	1190	97.7	233	8	ADP24510 PRO poly
19	1190	97.7	233	8	ADP23990 PRO poly
20	1190	97.7	249	8	ADQ97766 Human can
21	1187	97.5	233	8	ADQ80677 Rat anti-
22	1184	97.2	233	3	AA69969 Human Bcl
23	1184	97.2	233	3	AA83223 Bcl-x pol
24	1184	97.2	233	4	AAB50538 Human Bcl
25	1184	97.2	233	4	AA64262 Human Bcl

26	1184	97.2	233	4	AAB47515	Aab47515 Protein e
27	1184	97.2	233	7	ABW02410	ABW02410 Human bcl
28	1184	97.2	233	8	ADO19991	ADO19991 Human PRO
29	1184	97.2	233	8	ADQ80670	Adq80670 Human ant
30	1184	97.2	411	4	AAU00219	Asu00219 Bcl-XL-DT
31	1182	97.0	342	8	ADQ97763	Adq97763 Mouse can
32	1181	97.0	233	8	ADQ80679	Adq80679 Porcine a
33	1179	96.8	236	6	ABR83558	ABR83558 ToJA-BCL
34	1179	96.8	348	6	ABR83557	ABR83557 ToJA-BCL
35	1146.5	94.1	230	7	ADF60832	Adf60832 Human Bcl
36	1128	92.6	225	2	AAW19396	Aaw19396 "Deprenyl
37	1090	89.5	219	8	ADH52639	Adh52639 Chinese h
38	1079	88.6	219	8	ADH52641	Adh52641 Chinese h
39	1078	88.5	212	4	AAB20495	AB20495 Human Bcl
40	1078	88.5	212	8	ADG65210	ADG65210 Human Bcl
41	1075	88.3	212	4	AAG64285	Aag64285 Mutant bc
42	1061	87.1	485	4	AAU00222	Aau00222 LFn-Bcl-X
43	987.5	81.1	229	5	AAO18222	Aao18222 Human Bcl
44	979	80.4	235	2	AAW48312	Aaw48312 Mouse BCL
45	967	79.4	199	8	ADH52637	Adh52637 Chinese h

ALIGNMENTS

RESULT 1  
AAB73303  
ID AAB73303 standard; protein; 233 AA.  
XX  
AC AAB73303;  
XX  
DT 22-MAY-2001 (first entry)  
XX  
DE Rat wild-type Bcl-XL protein.  
XX  
KW Rat Bcl-XL; apoptosis inhibitor; programmed cell death inhibitor;  
KW wild-type; antiapoptotic; cell death-associated disease;  
KW tissue transplant preservative.  
XX  
OS Rattus norvegicus.  
XX  
PN WO200112807-A1.  
XX  
PD 22-FEB-2001.  
XX  
PF 17-AUG-2000; 2000WO-JP005502.  
XX  
PR 17-AUG-1999; 99JP-00230642.  
XX  
PI (NISC-) JAPAN SCI & TECHNOLOGY CORP.  
XX  
PI Ohta S, Asoh S;  
XX  
DR WPI; 2001-211219/21.  
DR N-ESDB; AAF75960.  
XX  
PT Modified cDNA of rat bcl-x gene and encoded protein with membrane permeability to enhance uptake for effective inhibition of cell death e.g. apoptosis, useful in remedies for diseases associated with cell death.

Claim 6; Page 45-46; 56pp; Japanese.  
The invention relates to a mutant rat Bcl-x protein and the cDNA encoding it. The mutant rat Bcl-x protein (Bcl-xFNK) has the substitutions Y22F, Q26N, and R165K relative to the wild-type Bcl-XL protein. The invention also encompasses recombinant vectors and host cells comprising the modified nucleic acid sequence. The mutant Bcl-x protein is able to permeate the cell membrane, thus enhancing its ability to be taken up into a cell and to act as an inhibitor of apoptosis (programmed cell death). Bcl-xFNK and nucleic acids encoding it are useful in remedies for diseases associated with cell death and in additives for maintaining the stability of transplanted cells and organs. The present sequence

```
CC represents wild-type rat Bcl-xL
XX
SQ Sequence 233 AA;

Query Match      100.0%; Score 1218; DB 4; Length 233;
Best Local Similarity 100.0%; Pred. No. 5.8e-121;
Matches 233; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSQSNRELVDVFLSYKLSQKGYSSQFSDVEENTEAPEETEPEPTPSAINGNPWHLA 60
Dy 1 MSQSNRELVDVFLSYKLSQKGYSSQFSDVEENTEAPEETEPEPTPSAINGNPWHLA 60
Qy 61 DSPAVNGATGHSSSLDAREVPMAAVKQALREAGDEFELRYRRAFSDLTSLQHTTPGTAY 120
Dy 61 DSPAVNGATGHSSSLDAREVPMAAVKQALREAGDEFELRYRRAFSDLTSLQHTTPGTAY 120
Qy 121 QSFEQVNNELFRDGVNNGRIVAFPSFGGALCVESVDKEMQVLVSRIASWMTYLNHLEP 180
Dy 121 QSFEQVNNELFRDGVNNGRIVAFPSFGGALCVESVDKEMQVLVSRIASWMTYLNHLEP 180
Qy 181 WIQENGWDTFVDLYGNNAASERKQGRFNRWFLTGMTVAGVVLGSLFSRK 233
Dy 181 WIQENGWDTFVDLYGNNAASERKQGRFNRWFLTGMTVAGVVLGSLFSRK 233

RESULT 2
ADE62921
ID ADE62921 standard; protein; 233 AA.
XX
AC ADE62921;
XX
DT 29-JAN-2004 (first entry)
XX
DE Rat Protein P53563, SEQ ID NO 8855.
XX
KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
XX chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX
OS Rattus norvegicus.
XX
PN WO2003016475-A2.
XX
PD 27-FEB-2003.
XX
PF 14-AUG-2002; 2002WO-US025765.
XX
PR 14-AUG-2001; 2001US-0312147P.
XX 01-NOV-2001; 2001US-0346382P.
XX 26-NOV-2001; 2001US-0333347P.
XX
PA (GEHO ) GEN HOSPITAL CORP.
XX
PI Woolf C, D'urso D, Befort K, Costigan M;
XX
PI WPI; 2003-268312/26.
XX
DR GENBANK; P53563.
XX
PT New composition comprising two or more isolated polypeptides, useful for
XX preparing a medicament for treating pain in an animal.
XX
PS Claim 1; Page; 1017pp; English.
XX
CC The invention discloses a composition comprising two or more isolated rat
XX or human polynucleotides or a polynucleotide which represents a fragment,
XX derivative or allelic variation of the nucleic acid sequence. Also
XX claimed are a vector comprising the novel polynucleotide, a host cell
XX comprising the vector, a method for identifying a nucleotide sequence
XX which is differentially regulated in an animal subjected to pain and a
XX kit to perform the method, an array, a method for identifying an agent
XX that increases or decreases the expression of the polynucleotide sequence
XX that is differentially expressed in neuronal tissue of a first animal
XX subjected to pain, a method for identifying a compound which regulates
```

```
CC the expression of a polynucleotide sequence which is differentially
XX expressed in an animal subjected to pain, a method for identifying a
XX compound that regulates the activity of one or more of the
XX polynucleotides, a method for producing a pharmaceutical composition, a
XX method for identifying a compound or small molecule that regulates the
XX activity in an animal of one or more of the polypeptides given in the
XX specification, a method for identifying a compound useful in treating
XX pain and a pharmaceutical composition comprising the one or more
XX polypeptides or their antibodies. The polynucleotide or the compound that
XX modulates its activity is useful for preparing a medicament for treating
XX pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
XX injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
XX therapy). The sequence presented is a rat protein (shown in Table 2 of
XX the specification) which is differentially expressed during pain. Note:
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic form directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.
SQ Sequence 233 AA;

Query Match      100.0%; Score 1218; DB 7; Length 233;
Best Local Similarity 100.0%; Pred. No. 5.8e-121;
Matches 233; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSQSNRELVDVFLSYKLSQKGYSSQFSDVEENTEAPEETEPEPTPSAINGNPWHLA 60
Dy 1 MSQSNRELVDVFLSYKLSQKGYSSQFSDVEENTEAPEETEPEPTPSAINGNPWHLA 60
Qy 61 DSPAVNGATGHSSSLDAREVPMAAVKQALREAGDEFELRYRRAFSDLTSLQHTTPGTAY 120
Dy 61 DSPAVNGATGHSSSLDAREVPMAAVKQALREAGDEFELRYRRAFSDLTSLQHTTPGTAY 120
Qy 121 QSFEQVNNELFRDGVNNGRIVAFPSFGGALCVESVDKEMQVLVSRIASWMTYLNHLEP 180
Dy 121 QSFEQVNNELFRDGVNNGRIVAFPSFGGALCVESVDKEMQVLVSRIASWMTYLNHLEP 180
Qy 181 WIQENGWDTFVDLYGNNAASERKQGRFNRWFLTGMTVAGVVLGSLFSRK 233
Dy 181 WIQENGWDTFVDLYGNNAASERKQGRFNRWFLTGMTVAGVVLGSLFSRK 233

RESULT 3
ADE62491
ID ADE62491 standard; protein; 233 AA.
XX
AC ADE62491;
XX
DT 29-JAN-2004 (first entry)
XX
DE Rat Protein P53563, SEQ ID NO 8420.
XX
KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
XX chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX
OS Rattus norvegicus.
XX
PN WO2003016475-A2.
XX
PD 27-FEB-2003.
XX
PF 14-AUG-2002; 2002WO-US025765.
XX
PR 14-AUG-2001; 2001US-0312147P.
XX 01-NOV-2001; 2001US-0346382P.
XX 26-NOV-2001; 2001US-0333347P.
XX
PA (GEHO ) GEN HOSPITAL CORP.
XX
PI Woolf C, D'urso D, Befort K, Costigan M;
XX
PI WPI; 2003-268312/26.
XX
DR GENBANK; P53563.
```



XX New composition comprising two or more isolated polypeptides, useful for  
PT preparing a medicament for treating pain in an animal.  
PS Claim 1; Page; 1017pp; English.  
XX  
XX The invention discloses a composition comprising two or more isolated rat  
CC or human polynucleotides or a polynucleotide which represents a fragment,  
CC derivative or allelic variation of the nucleic acid sequence. Also  
CC claimed are a vector comprising the novel polynucleotide, a host cell  
CC comprising the vector, a method for identifying a nucleotide sequence  
CC which is differentially regulated in an animal subjected to pain and a  
CC kit to perform the method, an array, a method for identifying an agent  
CC that increases or decreases the expression of the polynucleotide sequence  
CC that is differentially expressed in neuronal tissue of a first animal  
CC subjected to pain, a method for identifying a compound which regulates  
CC the expression of a polynucleotide sequence which is differentially  
CC expressed in an animal subjected to pain, a method for identifying a  
CC compound that regulates the activity of one or more of the  
CC polynucleotides, a method for producing a pharmaceutical composition, a  
CC method for identifying a compound or small molecule that regulates the  
CC activity in an animal of one or more of the polypeptides given in the  
CC specification, a method for identifying a compound useful in treating  
CC pain and a pharmaceutical composition comprising the one or more  
CC polypeptides or their antibodies. The polynucleotide or the compound that  
CC modulates its activity is useful for preparing a medicament for treating  
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
CC therapy). The sequence presented is a rat protein (shown in Table 2 of  
CC the specification) which is differentially expressed during pain. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic form directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 233 AA;

Query Match 100.0%; Score 1218; DB 7; Length 233;  
Best Local Similarity 100.0%; Pred. No. 5.8e-121;  
Matches 233; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MSQSNRELVDVFLSYKLSQKGYNSQFSQSDVEENRTEAPEETEPRETPSAINGNPSWHLA 60  
Db 1 MSQSNRELVDVFLSYKLSQKGYNSQFSQSDVEENRTEAPEETEPRETPSAINGNPSWHLA 60  
Qy 61 DSPAVNGATGHSSSLDAREVIPAQKALREAGDEPELRYRRFAFSDLTSLHTPTGTAY 120  
Db 61 DSPAVNGATGHSSSLDAREVIPAQKALREAGDEPELRYRRFAFSDLTSLHTPTGTAY 120  
Qy 121 QSPFQVNNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLVSRIASWATYLNHLEP 180  
Db 121 QSPFQVNNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLVSRIASWATYLNHLEP 180  
Qy 181 WIQENGWDTFVDLYGNNAAESKQGERFNRFLTGTAGVVLGSLFSRK 233  
Db 181 WIQENGWDTFVDLYGNNAAESKQGERFNRFLTGTAGVVLGSLFSRK 233

RESULT 4  
ADQ80678  
ID ADQ80678 standard; protein; 233 AA.  
XX  
XX ADQ80678;  
XX  
XX 21-OCT-2004 (first entry)  
XX Mouse anti-apoptotic Bcl-XL protein.  
XX Survival; neuron; tyrosine hydroxylase; tyrosine 3-monooxygenase; TH;  
XX anti-apoptotic; Bcl-XL; neurological disorder; neuroprotective;  
XX norepinephrine; antiparkinsonian; transplantation; drug screening;  
XX gene profiling; CNS disorder; neurodegenerative disease; mouse; murine.  
XX Mus musculus.

XX WO2004062554-A2.  
XX 29-JUL-2004.  
XX  
XX 07-JAN-2004; 2004WO-DK000008.  
XX  
XX 08-JAN-2003; 2003US-0438719P.  
XX 11-APR-2003; 2003DK-00000581.  
XX 22-APR-2003; 2003US-0464546P.  
XX (NSGE-) NSGENE AS.  
XX Martinez-Serrano A, Liste I, Villa A;  
XX WPI; 2004-544027/52.  
XX  
XX Enhancing the survival of neurons or cells expressing tyrosine  
PT hydroxylase (TH) for treating neurodegenerative disorders, comprises  
PT contacting neurons or TH expressing cells with Bcl-XL or its functional  
PT equivalent.  
XX  
XX Claim 21; SEQ ID NO 9; 108pp; English.  
XX  
XX The invention relates to a novel method for enhancing the survival of  
CC neurons and/or of cells expressing tyrosine hydroxylase (EC 1.14.16.2 -  
CC Tyrosine 3-monooxygenase) (TH + ). The method comprises contacting a  
CC population of cells with Bcl-XL or its functional equivalent, where the  
CC population of cells is selected from: neurons or cells capable of  
CC differentiating into neurons; or TH expressing cells or cells capable of  
CC differentiating into TH expressing cells. The invention further  
CC comprises: a composition of cells obtainable by the method above; a  
CC composition of isolated mammalian cells overexpressing the anti-apoptotic  
CC Bcl-XL protein; a neural progenitor cell; a differentiated dopaminergic  
CC neuron; an implantable cell culture device comprising: a semi-permeable  
CC membrane permitting the diffusion of a biologically active protein  
CC through it; and a composition of cells selected from above; a lentiviral  
CC vector particle being produced based on a lentiviral transfer vector;  
CC enhancing the survival of TH + cells in vivo; a retroviral particle being  
CC produced based on a retroviral transfer vector; enhancing the survival of  
CC in vivo differentiated dopaminergic neurons; a packaging cell line  
CC capable of producing an infective vector particle; a packaging cell line  
CC capable of producing an infective vector particle; treatment of a  
CC neurological disorder; a fusion protein comprising the Bcl-XL sequence  
CC comprising 233 amino acids ADQ80670 or its functional equivalent and a  
CC membrane translocation signal; an expression vector comprising a  
CC polynucleotide sequence coding for the fusion protein and a promoter  
CC sequence capable of directing the expression of the fusion protein in a  
CC host cell; a host cell comprising the expression vector; and producing  
CC the fusion protein. The compositions of the invention have  
CC neuroprotective, neurotropic, and antiparkinsonian activities. The cells  
CC are useful for transplantation, drug screening, gene profiling, or for  
CC the preparation of a medicament useful for the treatment of a CNS  
CC disorder. The CNS disorder is a neurodegenerative disease involving  
CC lesioned and traumatic neurons, including traumatic lesions of peripheral  
CC nerves, the medulla, the spinal cord, cerebral ischaemic neuronal  
CC damage, neuropathy, peripheral neuropathy, Alzheimer's disease,  
CC Huntington's disease, Parkinson's disease, amyotrophic lateral sclerosis,  
CC or memory impairment connected to dementia. The method is useful for  
CC enhancing the survival of neurons and/or of cells expressing tyrosine  
CC hydroxylase for the treatment of neurodegenerative disorders. This  
CC sequence represents a mouse anti-apoptotic Bcl-XL protein, used in the  
CC method for increasing the survival rate of neurons of the invention.  
XX  
XX Sequence 233 AA;

Query Match 99.3%; Score 1210; DB 8; Length 233;  
Best Local Similarity 99.6%; Pred. No. 4.1e-120;  
Matches 232; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 MSQSNRELVDVFLSYKLSQKGYNSQFSQSDVEENRTEAPEETEPRETPSAINGNPSWHLA 60  
Db 1 MSQSNRELVDVFLSYKLSQKGYNSQFSQSDVEENRTEAPEETEPRETPSAINGNPSWHLA 60

Qy 61 DSPAVNGATGSSSLDAREVIMAAVKQALREAGDEPELRYRRAFSDLTSLQHTTPGTAY 120  
 Db 61 DSPAVNGATGSSSLDAREVIMAAVKQALREAGDEPELRYRRAFSDLTSLQHTTPGTAY 120  
 Qy 121 QSFQVNVNELFRDGVNMGRIIVAFSFGALCVESVDKEMQVLVSRIASWMATYLNHLEP 180  
 Db 121 QSFQVNVNELFRDGVNMGRIIVAFSFGALCVESVDKEMQVLVSRIASWMATYLNHLEP 180  
 Qy 181 WIQENGWDFTVDLYGNNAASERKGOERFNRWFLTGMTVAGVVLGSLFSRK 233  
 Db 181 WIQENGWDFTVDLYGNNAASERKGOERFNRWFLTGMTVAGVVLGSLFSRK 233

RESULT 5  
 AAB73304  
 ID AAB73304 standard; protein; 233 AA.  
 AC AAB73304;  
 XX  
 DT 22-MAY-2001 (first entry)  
 DE  
 DE Mutant rat Bcl-xL protein, Bcl-xFNK.  
 XX  
 XX Rat Bcl-xL mutant; Bcl-xFNK; apoptosis inhibitor; membrane permeable;  
 KW programmed cell death inhibitor; wild-type; antiapoptotic;  
 KW cell death-associated disease; tissue transplant preservative; mutein.  
 XX  
 OS Rattus norvegicus.  
 OS Synthetic.  
 OS  
 XX WO200112807-A1.  
 XX  
 XX 22-FEB-2001.  
 XX  
 XX 17-AUG-2000; 2000WO-JP005502.  
 XX  
 XX 17-AUG-1999; 99JP-00230642.  
 XX  
 XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.  
 XX  
 XX Ohta S, Asoh S;  
 PI  
 XX WPI; 2001-211219/21.  
 XX  
 XX Modified cDNA of rat bcl-x gene and encoded protein with membrane  
 PT permeability to enhance uptake for effective inhibition of cell death  
 PT e.g. apoptosis, useful in remedies for diseases associated with cell  
 PT death.  
 XX  
 XX Claim 1; Page 46-47; 56pp; Japanese.  
 XX  
 XX The invention relates to a mutant rat Bcl-x protein and the cDNA encoding  
 CC it. The mutant rat Bcl-x protein (Bcl-xFNK) has the substitutions Y22F,  
 CC Q26N, and R165K relative to the wild-type Bcl-xL protein. The invention  
 CC also encompasses recombinant vectors and host cells comprising the  
 CC modified nucleic acid sequence. The mutant Bcl-x protein is able to  
 CC permeate the cell membrane, thus enhancing its ability to be taken up  
 CC into a cell and to act as an inhibitor of apoptosis (programmed cell  
 CC death). Bcl-xFNK and nucleic acids encoding it are useful in remedies for  
 CC diseases associated with cell death and in additives for maintaining the  
 CC stability of transplanted cells and organs. The present sequence  
 CC represents the mutant rat Bcl-xL protein, Bcl-xFNK  
 XX  
 SQ Sequence 233 AA;  
 Query Match 99.0%; Score 1206; DB 4; Length 233;  
 Best Local Similarity 98.7%; Pred. No. 1.1e-119;  
 Matches 230; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 61 DSPAVNGATGSSSLDAREVIMAAVKQALREAGDEPELRYRRAFSDLTSLQHTTPGTAY 120  
 Db 61 DSPAVNGATGSSSLDAREVIMAAVKQALREAGDEPELRYRRAFSDLTSLQHTTPGTAY 120  
 Qy 121 QSFQVNVNELFRDGVNMGRIIVAFSFGALCVESVDKEMQVLVSRIASWMATYLNHLEP 180  
 Db 121 QSFQVNVNELFRDGVNMGRIIVAFSFGALCVESVDKEMQVLVSRIASWMATYLNHLEP 180  
 Qy 181 WIQENGWDFTVDLYGNNAASERKGOERFNRWFLTGMTVAGVVLGSLFSRK 233  
 Db 181 WIQENGWDFTVDLYGNNAASERKGOERFNRWFLTGMTVAGVVLGSLFSRK 233

RESULT 6  
 ABG78480  
 ID ABG78480 standard; protein; 237 AA.  
 AC ABG78480;  
 XX  
 DT 15-NOV-2002 (first entry)  
 DE  
 DE Wild type BclX1 protein.  
 XX  
 XX Human; Bcl2; BclX1; programmed cell death; apoptosis; mutant; mutein.  
 XX Homo sapiens.  
 OS  
 XX WO200240530-A2.  
 XX  
 XX 23-MAY-2002.  
 XX  
 XX 15-NOV-2001; 2001WO-US045693.  
 XX  
 XX 20-NOV-2000; 2000US-00716395.  
 XX  
 XX (ABBO ) ABBOTT LAB.  
 XX  
 XX Pesik SW, Petros AM, Yoon H, Nettesheim DG;  
 XX WPI; 2002-490141/52.  
 XX  
 XX New mutant Bcl-2 proteins derived from a wild type human Bcl-2 protein,  
 PT useful in biological assays to identify substances that block the ability  
 PT of Bcl-2 to inhibit programmed cell death or apoptosis.  
 XX  
 XX Disclosure; Fig 1; 36pp; English.  
 XX  
 XX This invention relates to a novel mutant protein which is derived from a  
 CC wild type human Bcl-2 protein. The mutant is created by replacing a  
 CC sequence of amino acid residues comprising a flexible loop from the wild  
 CC type Bcl-2 protein with an amino acid sequence comprising at least two  
 CC acidic amino acids. The mutant Bcl-2 protein comprises a 166 residue  
 CC shown in the specification. The invention also comprises an assay for  
 CC identifying substances that bind to the Bcl-2 protein. The protein  
 CC sequences of the invention are useful in biological assays to identify  
 CC substances that block the ability of Bcl-2 to inhibit programmed cell  
 CC death or apoptosis. The present sequence represents a human Bcl2 mutant  
 CC protein used in the invention  
 XX  
 SQ Sequence 237 AA;  
 Query Match 98.3%; Score 1197; DB 5; Length 237;  
 Best Local Similarity 98.3%; Pred. No. 1e-118;  
 Matches 229; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MSQSNRELVDVFLSYKLSQKGYSMQSFSDVEENRTEAPEETEPEETPSAINGNPSWHLA 60  
 Db 5 MSQSNRELVDVFLSYKLSQKGYSMQSFSDVEENRTEAPEETEPEETPSAINGNPSWHLA 64  
 Qy 61 DSPAVNGATGSSSLDAREVIMAAVKQALREAGDEPELRYRRAFSDLTSLQHTTPGTAY 120  
 Db 65 DSPAVNGATGSSSLDAREVIMAAVKQALREAGDEPELRYRRAFSDLTSLQHTTPGTAY 124

Qy 1 MSQSNRELVDVFLSYKLSQKGYSMQSFSDVEENRTEAPEETEPEETPSAINGNPSWHLA 60  
 Db 5 MSQSNRELVDVFLSYKLSQKGYSMQSFSDVEENRTEAPEETEPEETPSAINGNPSWHLA 64  
 Qy 61 DSPAVNGATGSSSLDAREVIMAAVKQALREAGDEPELRYRRAFSDLTSLQHTTPGTAY 120  
 Db 65 DSPAVNGATGSSSLDAREVIMAAVKQALREAGDEPELRYRRAFSDLTSLQHTTPGTAY 124

QY 121 QSEFQVNNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLSRIASWMMATYLNHLEP 180  
DB 125 QSEFQVNNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLSRIASWMMATYLNHLEP 184

QY 181 WIQENGWDTFVDLYGNNAASRKQGERFNRWFLTGMTVAGVVLLGSLFSRK 233  
DB 185 WIQENGWDTFVELYGNNAASRKQGERFNRWFLTGMTVAGVVLLGSLFSRK 237

RESULT 7  
ADH52633  
ID ADH52633 standard; protein; 233 AA.  
XX AC ADH52633;  
XX 25-MAR-2004 (first entry)  
XX Chinese hamster anti-apoptosis bcl-xL wild-type protein.  
XX mammalian myeloma host cell; protein production; anti-apoptosis;  
XX cell death; Chinese hamster; bcl-xL; wild-type.  
XX Cricetulus griseus.  
XX US2003219871-A1.  
XX 27-NOV-2003.  
XX 28-MAR-2003; 2003US-00402017.  
XX 02-APR-2002; 2002US-0369307P.  
XX (BOEH ) BOEHRINGER INGELHEIM PHARMA GMBH & CO KG.  
XX Enkel B, Meents H, Fussenegger M;  
XX WPI; 2004-033642/03.  
XX N-PSDB; ADH52632.  
XX New genetically engineered hamster or murine myeloma host cells  
PT comprising enhanced levels of active anti-apoptosis genes, useful for  
PT producing complex protein therapeutics.  
XX Disclosure; SEQ ID NO 4; 46pp; English.

CC The invention relates to a novel mammalian host cell for producing  
CC protein therapeutics. The host cell comprises a hamster or a murine  
CC myeloma cell that is genetically modified by introduction of nucleic acid  
CC sequences encoding an anti-apoptosis gene, a selectable amplifiable  
CC marker gene and at least one gene of interest. The host cell of the  
CC invention may be useful for producing at least one protein encoded by a  
CC gene of interest. The DNA, polypeptide and the methods may be used for  
CC inhibiting or delaying cell death. The current sequence is that of the  
CC Chinese hamster anti-apoptosis bcl-xL wild-type protein of the invention.

XX SQ Sequence 233 AA;  
Query Match 98.2%; Score 1196; DB 8; Length 233;  
Best Local Similarity 98.3%; Pred. No. 1.3e-118;  
Matches 229; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSQSNRELVDFLSYKLSQKGYSSQSDVEENRTEAPEETPERETPSAINGNPSWHLA 60  
DB 1 MSQSNRELVDFLSYKLSQKGYSSQSDVEENRTEAPEETPERETPSAINGNPSWHLA 60

QY 61 DSPAVNGATGSHSSSLDAREVPMMAVKQALREAGDEFEFLRYRRAFSDLTSLHITPGTAY 120  
DB 61 DSPAVNGATGSHSSSLDAREVPMMAVKQALREAGDEFEFLRYRRAFSDLTSLHITPGTAY 120

QY 121 QSEFQVNNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLSRIASWMMATYLNHLEP 180  
DB 121 QSEFQVNNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLSRIASWMMATYLNHLEP 180

QY 181 WIQENGWDTFVDLYGNNAASRKQGERFNRWFLTGMTVAGVVLLGSLFSRK 233  
DB 181 WIQENGWDTFVELYGNNAASRKQGERFNRWFLTGMTVAGVVLLGSLFSRK 233

RESULT 8  
AAR68887  
ID AAR68887 standard; protein; 233 AA.  
XX AC AAR68887;  
XX 25-MAR-2003 (revised)  
DT 10-AUG-1995 (first entry)  
XX Human thymus BCL-XL.  
XX BCL-XL; apoptosis; cell death; cancer; neurodegenerative disease;  
KW autoimmune disease; Parkinson disease; amyotrophic lateral sclerosis;  
KW multiple sclerosis.  
XX Homo sapiens.  
XX WO9500642-A1.  
XX 05-JAN-1995.  
XX 22-JUN-1994; 94WO-US007089.  
XX 22-JUN-1993; 93US-00081448.  
XX (ARCH-) ARCH DEV CORP.  
XX (UNMI ) UNIV MICHIGAN.  
XX Thompson CB, Boise LH, Nunez G;  
XX WPI; 1995-052079/07.  
XX N-PSDB; AAQ81698.  
XX New poly-nucleotide encoding new poly-peptide(s) that modify apoptosis -  
PT and related vectors, recombinant cells and antibodies, useful in assay  
PT and for control of cell death in e.g. neuronal cells, lymphocytes and  
PT cancers.  
XX Claim 3; Page 94; 127pp; English.

CC This protein may be expressed recombinantly, particularly with pcMV  
CC plasmids as vectors for expression in mammalian cell cultures. The  
CC protein has particular application in cancer cells (failure of programmed  
CC cell death (PCD)) or neurodegenerative and autoimmune diseases (premature  
CC PCD), e.g. Parkinson's disease, amyotrophic lateral sclerosis and  
CC multiple sclerosis. (Updated on 25-MAR-2003 to correct FN field.)

XX SQ Sequence 233 AA;  
Query Match 97.7%; Score 1190; DB 2; Length 233;  
Best Local Similarity 97.9%; Pred. No. 5.6e-118;  
Matches 228; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MSQSNRELVDFLSYKLSQKGYSSQSDVEENRTEAPEETPERETPSAINGNPSWHLA 60  
DB 1 MSQSNRELVDFLSYKLSQKGYSSQSDVEENRTEAPEETPERETPSAINGNPSWHLA 60

QY 61 DSPAVNGATGSHSSSLDAREVPMMAVKQALREAGDEFEFLRYRRAFSDLTSLHITPGTAY 120  
DB 61 DSPAVNGATGSHSSSLDAREVPMMAVKQALREAGDEFEFLRYRRAFSDLTSLHITPGTAY 120

QY 121 QSEFQVNNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLSRIASWMMATYLNHLEP 180  
DB 121 QSEFQVNNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLSRIASWMMATYLNHLEP 180

QY 181 WIQENGWDTFVDLYGNNAASRKQGERFNRWFLTGMTVAGVVLLGSLFSRK 233  
DB 181 WIQENGWDTFVELYGNNAASRKQGERFNRWFLTGMTVAGVVLLGSLFSRK 233

Db	181	WQENGWDTFVELYGNNAAESRKQGERFNRWFLTGMTVAGVVLGSLFSRK	233
AAW05821			
ID	AAW05821	standard; protein; 233 AA.	
XX	AAW05821;		
XX	30-MAR-1997	(first entry)	
XX	Bcl-XL protein.		
XX	Human; bcl-XL; T-lymphocyte; cell death; BH1 domain; BH2 domain;		
KW	Bcl-2 homology domain; gene therapy; HIV; AIDS; antiseize;		
KW	immune disorder; autoimmune disease; graft rejection;		
KW	graft-versus-host disease; apoptosis; adoptive immunotherapy.		
XX	Homo sapiens.		
XX	Key	Location/Qualifiers	
FT	Domain	129..148	
FT	Domain	/note= "BH1 domain"	
FT	Domain	180..191	
XX	W09634956-A1.	/note= "BH2 domain"	
XX	07-NOV-1996.		
XX	02-MAY-1996;	96WO-US006203.	
XX	04-MAY-1995;	95US-00435518.	
PR	07-JUN-1995;	95US-00481739.	
XX	(USNA ) US SEC OF NAVY.		
PA	(ARCH-) ARCH DEV CORP.		
XX	June CH, Thompson CB;		
XX	WPI; 1996-506159/50.		
DR	N-ESDB; AAT40079.		
XX	Inducing or preventing death of T cells by bcl-XL protein regulation -		
PT	used to increase survival of HIV infected cells or to down:regulate		
PT	immune responses in immune diseases.		
XX	Disclosure; Page 52-53; 76pp; English.		
XX	This is the sequence of a human bcl-XL protein, which protects T-		
CC	lymphocytes against cell death. A splice variant form, bcl-XS, lacks a		
CC	stretch of 63 amino acids, and is a dominant negative regulator of bcl-XL		
CC	function. The gene may be modified to facilitate interaction with		
CC	costimulatory Bax protein and inhibit interaction with antagonistic Bad		
CC	protein, by modification of the Bcl-2 homology domains BH1 and/or BH2.		
CC	The bcl-XL gene may be introduced into T-cells in vivo or ex vivo via		
CC	gene transfer using a vector for HIV infection gene therapy, to augment		
CC	intracellular bcl-XL protein levels and protect from cell death. A		
CC	corresponding antisense oligonucleotide or expression vector may be used		
CC	in gene therapy of e.g. autoimmune disease, graft rejection or graft-		
CC	versus-host disease, to induce cell death (e.g. apoptosis) and down-		
CC	regulate the immune response in a T-lymphocyte population		
XX	Sequence 233 AA;		
XX	Query Match	97.7%; Score 1190; DB 2; Length 233;	
XX	Best Local Similarity	97.9%; Pred. No. 5.6e-118;	
XX	Matches 228; Conservative	2; Mismatches 3; Indels 0; Gaps 0;	
Qy	1	MSQSNRELVDVFLSKLSQKYSQSDVDVENRTEAPEETPERTPSAINGNPSWHLA	60
Db	1	MSQSNRELVDVFLSKLSQKYSQSDVDVENRTEAPEETPERTPSAINGNPSWHLA	60
Qy	61	DSPAVNGATGSHSSSLDAREVIPMAAVKQALREAGDEFELRYRRAFSDLTSLHTPGTAY	120
Db	61	DSPAVNGATGSHSSSLDAREVIPMAAVKQALREAGDEFELRYRRAFSDLTSLHTPGTAY	120
Qy	121	QSFEQVNNELFRDGVNMGRIIVAFPSFGALCVESVDKEMQVLVSRIASWMTYNDHLEP	180
Db	121	QSFEQVNNELFRDGVNMGRIIVAFPSFGALCVESVDKEMQVLVSRIASWMTYNDHLEP	180
Qy	181	WQENGWDTFVDLYGNNAAESRKQGERFNRWFLTGMTVAGVVLGSLFSRK	233
Db	181	WQENGWDTFVELYGNNAAESRKQGERFNRWFLTGMTVAGVVLGSLFSRK	233
XX	RESULT 10		
XX	AAW31530		
ID	AAW31530	standard; protein; 233 AA.	
XX	AAW31530;		
XX	19-FEB-1998	(first entry)	
XX	Human anti-apoptotic BCL-XL protein.		
XX	BCL-XL; anti-apoptotic protein; human; nuclear factor-kappa B;		
KW	NF-kappa B; inhibitor; organ transplant; tissue transplant; inflammation;		
KW	gene therapy; endothelial cell.		
XX	Homo sapiens.		
XX	W09730083-A1.		
XX	21-AUG-1997.		
XX	13-FEB-1997;	97WO-EP000676.	
XX	14-FEB-1996;	96US-00601515.	
PR	19-APR-1996;	96US-00634995.	
XX	(NOVS ) NOVARTIS AG.		
PA	(NEWEE-) NEW ENGLAND DEACONESS HOSPITAL.		
XX	Bach FH, Ferran C;		
XX	WPI; 1997-424975/39.		
XX	Recombinant endothelial cell containing DNA encoding anti-apoptotic		
PT	protein - is less susceptible to inflammatory response and is useful for		
PT	generating tissues or organs for transplantation.		
XX	Claim 6; Page 46; 75pp; English.		
XX	This protein sequence comprises human BCL-XL, a protein capable of		
CC	blocking or suppressing NF-kappa B (NF-KB) activation. A claimed method		
CC	of genetically modifying a mammalian endothelial cell to render it less		
CC	susceptible to an inflammatory or other immunological stimulus comprises		
CC	inserting into the cell, DNA encoding an anti-apoptotic protein able to		
CC	inhibit NF-KB, and expressing the cell such that NF-KB activation of the		
CC	cell is inhibited in the presence of the cellular activating		
CC	stimulus. Suitable anti-apoptotic proteins include A20 (see AAW31528),		
CC	BCL-2 (see AAW31529), BCL-XL and A1 (see AAW31531) and their deletion		
CC	mutants capable of inhibiting NF-KB, such as polypeptides comprising		
CC	amino acid residues 5-24, 86-100, 129-148 and 180-195 of BCL-XL. Also		
CC	claimed are: (1) a mammalian endothelial cell modified by the above		
CC	method; and (2) a non-human transgenic or somatic recombinant mammal		
CC	comprising DNA encoding an anti-apoptotic protein of a different species.		
CC	The method can be used to generate donor endothelial cells or graftable		
CC	tissues or organs for transplantation into recipient species		
XX	Sequence 233 AA;		
XX	Query Match	97.7%; Score 1190; DB 2; Length 233;	
XX	Best Local Similarity	97.9%; Pred. No. 5.6e-118;	
XX	Matches 228; Conservative	2; Mismatches 3; Indels 0; Gaps 0;	

Qy 1 MSQSNRELVDVFLSYKLSQKGYWSQSDVEENRTEAPEETEPEPESAINGNPSWHLA 60  
 Db 1 MSQSNRELVDVFLSYKLSQKGYWSQSDVEENRTEAPEETEPEPESAINGNPSWHLA 60  
 Qy 61 DSPAVNGATGHSSSLDAREVPMVAVKQALREAGDEFELRYRRAFSDLTSLQHTTPGTAY 120  
 Db 61 DSPAVNGATGHSSSLDAREVPMVAVKQALREAGDEFELRYRRAFSDLTSLQHTTPGTAY 120  
 Qy 121 QSFQVNVNELFRDGVNMGRIVAFPSFGALCVESVDKEMQVLVSRIASWMTATYLNHLEP 180  
 Db 121 QSFQVNVNELFRDGVNMGRIVAFPSFGALCVESVDKEMQVLVSRIASWMTATYLNHLEP 180  
 Qy 181 WIQENGCGWDTFVLYGNNAASRKQGERFNRWFLTGMTVAGVVLGSLFSRK 233  
 Db 181 WIQENGCGWDTFVLYGNNAASRKQGERFNRWFLTGMTVAGVVLGSLFSRK 233

RESULT 11  
 ADE62493  
 ID ADE62493 standard; protein; 233 AA.  
 XX AC ADE62493;  
 XX DT 29-JAN-2004 (first entry)  
 XX DE Human Protein Q07817, SEQ ID NO 8422.  
 XX KW Human; pain; neuronal tissue; gene therapy;  
 XX KW spinal segmental nerve injury; chronic constriction injury; CCI;  
 XX KW spared nerve injury; SNI; Chung.  
 XX OS Homo sapiens.  
 XX PN WO2003016475-A2.  
 XX PD 27-FEB-2003.  
 XX PF 14-AUG-2002; 2002WO-US025765.  
 XX PR 14-AUG-2001; 2001US-0312147P.  
 XX PR 01-NOV-2001; 2001US-0346382P.  
 XX PR 26-NOV-2001; 2001US-0333347P.  
 XX PA (GEHO ) GEN HOSPITAL CORP.  
 XX PA (FARB ) BAYER AG.  
 XX PI Woolf C, D'urso D, Befort K, Costigan M;  
 XX WPI; 2003-268312/26.  
 XX DR GENBANK; Q07817.  
 XX PT New composition comprising two or more isolated polypeptides, useful for  
 XX PT preparing a medicament for treating pain in an animal.  
 XX PS Claim 1; Page; 101pp; English.

XX The invention discloses a composition comprising two or more isolated rat  
 CC or human polynucleotides or a polynucleotide which represents a fragment,  
 CC derivative or allelic variation of the nucleic acid sequence. Also  
 CC claimed are a vector comprising the novel polynucleotide, a host cell  
 CC comprising the vector, a method for identifying a nucleotide sequence  
 CC which is differentially regulated in an animal subjected to pain and a  
 CC kit to perform the method, an array, a method for identifying an agent  
 CC that increases or decreases the expression of the polynucleotide sequence  
 CC that is differentially expressed in neuronal tissue of a first animal  
 CC subjected to pain, a method for identifying a compound which regulates  
 CC the expression of a polynucleotide sequence which is differentially  
 CC expressed in an animal subjected to pain, a method for identifying a  
 CC compound that regulates the activity of one or more of the  
 CC polynucleotides, a method for producing a pharmaceutical composition, a  
 CC method for identifying a compound or small molecule that regulates the  
 CC activity in an animal of one or more of the polypeptides given in the

CC specification, a method for identifying a compound useful in treating  
 CC pain and a pharmaceutical composition comprising the one or more  
 CC polypeptides or their antibodies. The polynucleotide or the compound that  
 CC modulates its activity is useful for preparing a medicament for treating  
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
 CC therapy). The sequence presented is a human protein (shown in Table 2 of  
 CC the specification) which is differentially expressed during pain. Note:  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic form directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 XX Sequence 233 AA;

Query Match 97.7%; Score 1190; DB 7; Length 233;  
 Best Local Similarity 97.9%; Pred. No. 5.6e-118;  
 Matches 228; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 Qy 1 MSQSNRELVDVFLSYKLSQKGYWSQSDVEENRTEAPEETEPEPESAINGNPSWHLA 60  
 Db 1 MSQSNRELVDVFLSYKLSQKGYWSQSDVEENRTEAPEETEPEPESAINGNPSWHLA 60  
 Qy 61 DSPAVNGATGHSSSLDAREVPMVAVKQALREAGDEFELRYRRAFSDLTSLQHTTPGTAY 120  
 Db 61 DSPAVNGATGHSSSLDAREVPMVAVKQALREAGDEFELRYRRAFSDLTSLQHTTPGTAY 120  
 Qy 121 QSFQVNVNELFRDGVNMGRIVAFPSFGALCVESVDKEMQVLVSRIASWMTATYLNHLEP 180  
 Db 121 QSFQVNVNELFRDGVNMGRIVAFPSFGALCVESVDKEMQVLVSRIASWMTATYLNHLEP 180  
 Qy 181 WIQENGCGWDTFVLYGNNAASRKQGERFNRWFLTGMTVAGVVLGSLFSRK 233  
 Db 181 WIQENGCGWDTFVLYGNNAASRKQGERFNRWFLTGMTVAGVVLGSLFSRK 233

RESULT 12  
 ADM45995  
 ID ADM45995 standard; protein; 233 AA.  
 XX AC ADM45995;  
 XX DT 03-JUN-2004 (first entry)  
 XX DE Human apoptosis inhibitory factor Bcl-xL protein.  
 XX KW random oligonucleotide library; protein interaction; ligand;  
 XX KW receptor binding site; human; apoptosis inhibitory factor; Bcl-xL.  
 XX OS Homo sapiens.  
 XX PN JP2004024078-A.  
 XX PD 29-JAN-2004.  
 XX PF 24-JUN-2002; 2002JP-00183456.  
 XX PR 24-JUN-2002; 2002JP-00183456.  
 XX PA (SRE-) SERESUTA REKISHIKO SCI KK.  
 XX WPI; 2004-161478/16.  
 XX DR N-PSDB; ADM45994.

XX Random oligonucleotide useful for detecting protein interaction, having  
 PT base sequence, where each base of 1st and 2nd of the codon is the any of  
 PT G, C, T (U), or A and the base of 3rd of codon is G or C, or G or T (U).  
 XX Example 2; SEQ ID NO 4; 43pp; Japanese.  
 XX The invention relates to a novel random oligonucleotide having a base  
 CC sequence where each base of the 1st and 2nd codon is any of G, C, T (U)  
 CC or A and each base of the 3rd codon is G or C, or G or T (U). The methods  
 CC of the invention may be useful for preparing a random oligonucleotide

CC preparation to be used for detecting protein interactions or for  
 CC screening ligand or receptor protein binding sites. The current sequence  
 CC is that of the human apoptosis inhibitory factor Bcl-xL protein of the  
 CC invention.

XX SQ Sequence 233 AA;

Query Match 97.7%; Score 1190; DB 8; Length 233;  
 Best Local Similarity 97.9%; Pred. No. 5.6e-118;  
 Matches 228; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MSQSNRELVDVFLSYKLSQKGYNSQFSDVEENRTEAPEETEPEPTPSAINGNPSWHLA 60  
 |||||  
 Db 1 MSQSNRELVDVFLSYKLSQKGYNSQFSDVEENRTEAPEETEPEPTPSAINGNPSWHLA 60  
 |||||

Qy 61 DSPAVNGATGHSSSLDAREVTPMAAVKQALREAGDEPELRYRRAPFSDLTSQLHTTPGTAY 120  
 |||||  
 Db 61 DSPAVNGATGHSSSLDAREVTPMAAVKQALREAGDEPELRYRRAPFSDLTSQLHTTPGTAY 120  
 |||||

Qy 121 QSFQVNVNELFRDGVNMGRIIVAFPSFGALCVESVDKEMQVLVSRIASWMTATYLNHLEP 180  
 |||||  
 Db 121 QSFQVNVNELFRDGVNMGRIIVAFPSFGALCVESVDKEMQVLVSRIASWMTATYLNHLEP 180  
 |||||

Qy 181 WIQENGWDTFVDLYGNNAAESRKQERFNRWFLTGMTVAGVVLGSLFSRK 233  
 |||||  
 Db 181 WIQENGWDTFVELYGNNAAESRKQERFNRWFLTGMTVAGVVLGSLFSRK 233  
 |||||

RESULT 13  
 ADN04261  
 ID ADN04261 standard; protein; 233 AA.  
 XX  
 AC ADN04261;  
 XX  
 DT 01-JUL-2004 (first entry)  
 XX  
 DE Antipsoriatic protein sequence #325.  
 XX  
 KW antipsoriatic; gene therapy; psoriasis; diagnosis.  
 XX  
 OS Homo sapiens.  
 XX  
 FN WO2004028479-A2.  
 XX  
 PD 08-APR-2004.  
 XX  
 XX 25-SEP-2003; 2003WO-US030907.  
 XX  
 XX 25-SEP-2002; 2002US-0414006P.  
 XX  
 PA (GETH ) GENENTECH INC.  
 XX  
 XX Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;  
 PI Wu TD;  
 XX  
 XX WPI: 2004-305105/28.  
 DR  
 DR N-PSDB; ADN04260.  
 XX  
 PT New PRO nucleic acid or polypeptide, useful for preparing a  
 PT pharmaceutical composition for diagnosing or treating psoriasis in a  
 PT mammal.  
 XX  
 XX Claim 9; SEQ ID NO 655; 3069pp; English.  
 FS  
 XX The invention relates to novel polynucleotide and polypeptides for  
 CC treating psoriasis or a sequence having at least 80% identity to the  
 CC above sequences. The nucleic acid is useful for preparing a composition  
 CC for diagnosing or treating psoriasis in a mammal. This sequence  
 CC corresponds to one of the polypeptides of the invention.  
 XX  
 SQ Sequence 233 AA;

Query Match 97.7%; Score 1190; DB 8; Length 233;  
 Best Local Similarity 97.9%; Pred. No. 5.6e-118;  
 Matches 228; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MSQSNRELVDVFLSYKLSQKGYNSQFSDVEENRTEAPEETEPEPTPSAINGNPSWHLA 60  
 |||||  
 Db 1 MSQSNRELVDVFLSYKLSQKGYNSQFSDVEENRTEAPEETEPEPTPSAINGNPSWHLA 60  
 |||||

Qy 61 DSPAVNGATGHSSSLDAREVTPMAAVKQALREAGDEPELRYRRAPFSDLTSQLHTTPGTAY 120  
 |||||  
 Db 61 DSPAVNGATGHSSSLDAREVTPMAAVKQALREAGDEPELRYRRAPFSDLTSQLHTTPGTAY 120  
 |||||

Qy 121 QSFQVNVNELFRDGVNMGRIIVAFPSFGALCVESVDKEMQVLVSRIASWMTATYLNHLEP 180  
 |||||  
 Db 121 QSFQVNVNELFRDGVNMGRIIVAFPSFGALCVESVDKEMQVLVSRIASWMTATYLNHLEP 180  
 |||||

Qy 181 WIQENGWDTFVDLYGNNAAESRKQERFNRWFLTGMTVAGVVLGSLFSRK 233  
 |||||  
 Db 181 WIQENGWDTFVELYGNNAAESRKQERFNRWFLTGMTVAGVVLGSLFSRK 233  
 |||||

RESULT 14  
 ADN04261  
 ID ADN04261 standard; protein; 233 AA.  
 XX  
 AC ADN04261;  
 XX  
 DT 12-AUG-2004 (first entry)  
 XX  
 DE Human PRO polypeptide #394.  
 XX  
 KW Human; PRO; immune related disorder; systemic lupus erythematosus;  
 KW rheumatoid arthritis; osteoarthritis; juvenile chronic arthritis;  
 KW systemic sclerosis; Sjogren's syndrome; vasculitis; sarcoidosis;  
 KW autoimmune haemolytic anaemia; autoimmune thrombocytopenia; thyroiditis;  
 KW diabetes mellitus; renal disease; demyelinating disease;  
 KW central nervous system; peripheral nervous system;  
 KW demyelinating polyneuropathy; Guillain-Barre syndrome;  
 KW chronic inflammatory demyelinating polyneuropathy.  
 XX  
 OS Homo sapiens.  
 XX  
 FN WO2004043361-A2.  
 XX  
 PD 27-MAY-2004.  
 XX  
 XX 06-NOV-2003; 2003WO-US035268.  
 PF  
 XX 08-NOV-2002; 2002US-0425235P.  
 PR  
 XX (GETH ) GENENTECH INC.  
 PA  
 XX Fong S, Dennis K, Clark H, Chiu H, Schoenfeld J, Williams PM;  
 PI Wood WI, Wu TD;  
 XX  
 XX WPI: 2004-420067/39.  
 DR  
 DR N-PSDB; ADN04260.  
 XX  
 PT Novel PRO polypeptide e.g., PRO69614, PRO71106, or PRO86388 useful for  
 PT treating an immune related disorder such as systemic lupus erythematosus,  
 PT rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis or  
 PT spondyloarthropathy.  
 XX  
 XX Claim 7; SEQ ID NO 788; 1731pp; English.  
 FS  
 XX The invention relates to human PRO polypeptides and the polynucleotides  
 CC encoding them. The polypeptides and polynucleotides are useful for  
 CC treating and diagnosing immune related disorders in mammals. The immune  
 CC related disorders include systemic lupus erythematosus, rheumatoid  
 CC arthritis, osteoarthritis, juvenile chronic arthritis, systemic  
 CC sclerosis, Sjogren's syndrome, vasculitis, sarcoidosis, autoimmune  
 CC haemolytic anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes  
 CC mellitus, immune-mediated renal disease, demyelinating diseases of the  
 CC central or peripheral nervous system, demyelinating polyneuropathy.

CC Guillain-Barre syndrome and chronic inflammatory demyelinating  
CC polyneuropathy. This sequence represents a human PRO polypeptide of the  
CC invention.  
XX  
SQ Sequence 233 AA;  
Query Match 97.7%; Score 1190; DB 8; Length 233;  
Best Local Similarity 97.9%; Pred. No. 5.6e-118;  
Matches 228; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
Qy 1 MSQSNRELVDVFLSYKLSQKGYSSQFSDVEENTEAPEETEPETPSAINGNPSWHLA 60  
Db 1 MSQSNRELVDVFLSYKLSQKGYSSQFSDVEENTEAPEETEPETPSAINGNPSWHLA 60  
Qy 61 DSPAVNGATGHSSSLDAREVPMVAVKQALREAGDEPELRYRRFSDLTSLHITPGTAY 120  
Db 61 DSPAVNGATGHSSSLDAREVPMVAVKQALREAGDEPELRYRRFSDLTSLHITPGTAY 120  
Qy 121 QSPQVNVNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLVSRIASWMTYLNHLEP 180  
Db 121 QSPQVNVNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLVSRIASWMTYLNHLEP 180  
Qy 181 WIQENGWDTFVDLYGNNAAESKQGERFNRWFLTGMTVAGVVLGSLFSRK 233  
Db 181 WIQENGWDTFVDLYGNNAAESKQGERFNRWFLTGMTVAGVVLGSLFSRK 233  
RESULT 15  
ABM82217  
ID ABM82217 standard; protein; 233 AA.  
AC  
XX  
XX  
DT 18-NOV-2004 (first entry)  
XX  
XX  
DE Tumour-associated antigenic target (TAT) polypeptide PRO83141, SEQ.5714.  
XX  
KW Tumour-associated antigenic target; TAT; human; overexpression; cancer;  
KW tumour; diagnosis; cell proliferative disorder; breast cancer;  
KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;  
KW central nervous system cancer; bladder cancer; pancreatic cancer;  
KW cervical cancer; melanoma; leukaemia; hybridisation probe;  
KW chromosome identification; chromosome mapping; gene mapping;  
KW gene therapy; cytostatic.  
XX  
OS Homo sapiens.  
XX  
XX WO2004030615-A2.  
XX  
XX 15-APR-2004.  
XX  
XX 29-SEP-2003; 2003WO-US028547.  
XX  
XX 02-OCT-2002; 2002US-0414971P.  
XX  
XX (GETH ) GENENTECH INC.  
XX  
XX Wu TD, Zhang Z, Zhou Y;  
XX  
XX WPI; 2004-347921/32.  
XX  
XX N-PSDB; ACN40740.  
XX  
XX New tumor-associated antigenic target polypeptides and nucleic acids,  
XX useful in preparing a medicament for treating or detecting a  
XX proliferative disorder, e.g. breast, lung, colorectal, ovarian or  
XX prostate cancer or tumor.  
XX  
XX Claim 12; SEQ ID NO 5714; 7273pp; English.  
XX  
XX The invention relates to human tumour-associated antigenic target (TAT)  
XX polypeptides, and their related nucleic acids. The TAT polypeptides are  
XX overexpressed in cancer tissues compared to normal tissues, and may thus  
XX serve as effective targets for the diagnosis and treatment of cancer in

CC mammals. The invention also relates to nucleic acid and polypeptide  
CC sequences at least 80% identical to the TAT nucleic acids and  
CC polypeptides; expression vectors and host cells comprising a TAT nucleic  
CC acid; an antibody specific for a TAT polypeptide; a peptide or organic  
CC molecule which binds to a TAT polypeptide; fusion proteins comprising a  
CC TAT polypeptide; and methods and compositions for the treatment or  
CC diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,  
CC antibodies, antagonists, binding molecules and compositions are useful  
CC for diagnosing or treating a cell proliferative disorder associated with  
CC increased TAT expression, particularly cancers such as breast cancer,  
CC colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder  
CC cancer, pancreatic cancer, cervical cancer, cancers of the central  
CC nervous system, melanoma and leukaemia. TAT nucleic acids may further be  
CC used as hybridisation probes, in chromosome and gene mapping, in  
CC chromosome identification and in gene therapy. The present sequence  
CC represents a TAT polypeptide of the invention  
XX  
SQ Sequence 233 AA;  
Query Match 97.7%; Score 1190; DB 8; Length 233;  
Best Local Similarity 97.9%; Pred. No. 5.6e-118;  
Matches 228; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
Qy 1 MSQSNRELVDVFLSYKLSQKGYSSQFSDVEENTEAPEETEPETPSAINGNPSWHLA 60  
Db 1 MSQSNRELVDVFLSYKLSQKGYSSQFSDVEENTEAPEETEPETPSAINGNPSWHLA 60  
Qy 61 DSPAVNGATGHSSSLDAREVPMVAVKQALREAGDEPELRYRRFSDLTSLHITPGTAY 120  
Db 61 DSPAVNGATGHSSSLDAREVPMVAVKQALREAGDEPELRYRRFSDLTSLHITPGTAY 120  
Qy 121 QSPQVNVNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLVSRIASWMTYLNHLEP 180  
Db 121 QSPQVNVNELFRDGVNNGRIVAFPSFGALCVESVDKEMQVLVSRIASWMTYLNHLEP 180  
Qy 181 WIQENGWDTFVDLYGNNAAESKQGERFNRWFLTGMTVAGVVLGSLFSRK 233  
Db 181 WIQENGWDTFVDLYGNNAAESKQGERFNRWFLTGMTVAGVVLGSLFSRK 233  
Search completed: March 31, 2005, 00:28:16  
Job time : 169.1 secs

100



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 31, 2005, 00:31:02 ; Search time 7391 Seconds  
(without alignments)  
11420.510 Million cell updates/sec

Title: US-10-049-822A-1  
Perfect score: 1742  
Sequence: 1 cacagacagaccagctgag.....ggagccctggggcttcct 1742

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl.\*

1: gb\_ba.\*

2: gb\_hcg.\*

3: gb\_in.\*

4: gb\_on.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1742	100.0	1742	6	BD012974 A mutagenized rat bcl-x cDNA and a modified protein therefrom
2	1742	100.0	1742	6	BD013799 Modified
3	1742	100.0	1742	10	U72350 Rattus norv
4	1662.2	95.4	2232	10	X82537 R.norvegicus
5	1240.6	71.2	172803	2	AC147787 Rattus no
6	1240.6	71.2	227938	2	AC098008 Rattus no
7	1168	67.0	1466	6	AX775078 Sequence
8	1168	67.0	1466	10	MMU51278
9	1085	62.3	224136	2	AC133260 Rattus no
10	1085	62.3	249872	2	AC106107 Rattus no
11	1038.2	59.6	2575	6	CQ827863 Sequence
12	1038.2	59.6	2575	9	BC019307 Homo sapi
13	893.4	51.3	1235	10	MMU10102
14	811.6	46.6	8022	10	AF133277S3
15	811.6	46.6	157220	10	AL731857 Mouse DNA
16	792.2	45.5	1684	10	MMBCLXP2
17	787.8	45.2	863	6	AX925688 Sequence
18	769.6	44.2	979	10	MUSBCLX
19	754.6	43.3	764	10	RNU10579 Rattus norv

20	741	42.5	926	6	AR118504 Sequence
21	741	42.5	926	6	CQ765842 Sequence
22	737.8	42.4	926	6	AR054021 Sequence
23	737.8	42.4	926	6	AR124952 Sequence
24	737.8	42.4	926	6	AR144311 Sequence
25	737.8	42.4	926	6	AR172594 Sequence
26	737.8	42.4	926	6	BD243042 Antisense
27	737.8	42.4	926	6	E58777 Screening m
28	737.8	42.4	926	6	IS2011 Sequence 5
29	737.8	42.4	926	6	AR380885 Sequence
30	737.8	42.4	926	6	AR559339 Sequence
31	737.8	42.4	926	6	AX839772 Sequence
32	737.8	42.4	926	6	AX925686 Sequence
33	737.8	42.4	926	9	HSBCLXL
34	736.2	42.3	926	6	AR371661 Sequence
35	721.8	41.4	5903	6	CQ879092 Sequence
36	721.8	41.4	6535	6	CQ879090 Sequence
37	721.8	41.4	7257	6	CQ879091 Sequence
38	714.8	41.0	726	10	U34963 Rattus norv
39	714.8	41.0	726	10	S76513 bcl-x-apopt
40	711.4	40.8	1252	4	AB073983 Canis fam
41	703.6	40.4	7436	6	CQ879093 Sequence
42	690.8	39.7	702	6	CQ879094 Sequence
43	690	39.6	1163	4	AB080951 Felis cat
44	671.8	38.6	766	4	AF164517 Ovis arie
45	665.2	38.2	702	10	MMU10101 Mus muscula

ALIGNMENTS

RESULT 1  
BD012974

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BD012974 1742 bp DNA linear PAT 02-AUG-2002  
A mutagenized rat bcl-x cDNA and a modified protein therefrom.  
BD012974  
BD012974.1 GI:22093163  
WO 0112807-A/1.  
Rattus norvegicus (Norway rat)  
Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
1 (bases 1 to 1742)  
Ota,S. and Aso,S.  
A mutagenized rat bcl-x cDNA and a modified protein therefrom  
Patent: WO 0112807-A 1 22-FEB-2001;  
JAPAN SCIENCE AND TECHNOLOGY CORP, SHIGEO OTA, SADAMITSU ASO  
OS Rattus norvegicus (rat)  
PN WO 0112807-A/1  
PD 22-FEB-2001  
PF 17-AUG-2000 WO 2000JP005502  
PR 17-AUG-1999 JP 99P 230642  
PI SHIGEO OTA, SADAMITSU ASO  
PC C12N15/12,C12N5/06,C07K14/47  
CC CC  
FH Key Location/Qualifiers  
FT CDS  
1..1742 Location/Qualifiers  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"

Query Match 100.0%; Score 1742; DB 6; Length 1742;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1742; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CACAGACAGACCCAGTGAGTGGACGAGTGTGTTGGACAAATGGACTGGTGGAGCCCATCT 60

Db 1 CACAGACAGACCCAGTGAGTGGACGAGTGTGTTGGACAAATGGACTGGTGGAGCCCATCT 60

QY 61 CTATTATATAAATGTCTCAGAGCAACCGGAGCTGGTGTGCTGACTTTCTCTCTACAAAG 120  
DB |||||  
QY 61 CTATTATATAAATGTCTCAGAGCAACCGGAGCTGGTGTGCTGACTTTCTCTCTACAAAG 120  
DB |||||  
QY 121 TCTCCAGAAAGGATACAGCTGGAGTCAAGTTAGCGATGTGAGAGAAACAGGACTGAAG 180  
DB |||||  
QY 121 TCTCCAGAAAGGATACAGCTGGAGTCAAGTTAGCGATGTGAGAGAAACAGGACTGAAG 180  
DB |||||  
QY 181 CCCAGAGAAACCTGAACCCAGAAAGGAGAGCCCGAGTGCCATCAATGGCAACCCATCCT 240  
DB |||||  
QY 181 CCCAGAGAAACCTGAACCCAGAAAGGAGAGCCCGAGTGCCATCAATGGCAACCCATCCT 240  
DB |||||  
QY 241 GGCACCTGGCGATAGCCCGCGTGAATGAGGCCACTGGCCACAGCAGCAGTTGGATG 300  
DB |||||  
QY 241 GGCACCTGGCGATAGCCCGCGTGAATGAGGCCACTGGCCACAGCAGCAGTTGGATG 300  
DB |||||  
QY 301 CGCGGAGGATATCCCCATGGCAGCAGTGAAGCAAGCGCTGAGAGAGCTGGCGATGAGT 360  
DB |||||  
QY 301 CGCGGAGGATATCCCCATGGCAGCAGTGAAGCAAGCGCTGAGAGAGCTGGCGATGAGT 360  
DB |||||  
QY 361 TTGAATCGGGTACCGGAGAGCAATTCAGTGATCTAAATCCAGCTTCATATACCCAG 420  
DB |||||  
QY 361 TTGAATCGGGTACCGGAGAGCAATTCAGTGATCTAAATCCAGCTTCATATACCCAG 420  
DB |||||  
QY 421 GGACAGCATATCAGAGCTTTGAACAGGTAGTGAATGAACCTCTTTCCGGATGGGGTAAACT 480  
DB |||||  
QY 421 GGACAGCATATCAGAGCTTTGAACAGGTAGTGAATGAACCTCTTTCCGGATGGGGTAAACT 480  
DB |||||  
QY 481 GGGGTGCGATGTGGCCCTCTCTCTTTGGCGGGCACTGTGCGTGAAGCGGTAGACA 540  
DB |||||  
QY 481 GGGGTGCGATGTGGCCCTCTCTCTTTGGCGGGCACTGTGCGTGAAGCGGTAGACA 540  
DB |||||  
QY 541 AGGAGATCGAGTATTTGGTGTAGTGGATGCAAGTTGGATGGCCACCTACCTGATGACC 600  
DB |||||  
QY 541 AGGAGATCGAGTATTTGGTGTAGTGGATGCAAGTTGGATGGCCACCTACCTGATGACC 600  
DB |||||  
QY 601 ACCTAGAGCTTGGATCCAGAGAAACCGCGGTGGGACACTTTTGTGATCTCTACGGGA 660  
DB |||||  
QY 601 ACCTAGAGCTTGGATCCAGAGAAACCGCGGTGGGACACTTTTGTGATCTCTACGGGA 660  
DB |||||  
QY 661 ACAATGCAAGCCGAGAGCCGGAAAGCCAGGAGCTTTCAACCGCTGGTTCTTGACGG 720  
DB |||||  
QY 661 ACAATGCAAGCCGAGAGCCGGAAAGCCAGGAGCTTTCAACCGCTGGTTCTTGACGG 720  
DB |||||  
QY 721 GCATGACTGTGGCTGGTGTAGTTCTGTGGGCTCACTCTTCAGTGGGAAGTGACAGACA 780  
DB |||||  
QY 721 GCATGACTGTGGCTGGTGTAGTTCTGTGGGCTCACTCTTCAGTGGGAAGTGACAGACA 780  
DB |||||  
QY 781 CTGACCGTCCACTCACTCACTCCACCTCCACCTTGCCCCCACCACCAACTCTCTTTACGCC 840  
DB |||||  
QY 781 CTGACCGTCCACTCACTCACTCCACCTCCACCTTGCCCCCACCACCAACTCTCTTTACGCC 840  
DB |||||  
QY 841 ACCATTGTCTACAGGAGAACCACTACATGCAACTCAGCCCTTCCCTTATTATAGGGTT 900  
DB |||||  
QY 841 ACCATTGTCTACAGGAGAACCACTACATGCAACTCAGCCCTTCCCTTATTATAGGGTT 900  
DB |||||  
QY 901 GGGCTAGACGAGTCCCTGAGTGTAGTGTCTAGAAATCTACACGCTTCTGTGAAGC 960  
DB |||||  
QY 901 GGGCTAGACGAGTCCCTGAGTGTAGTGTCTAGAAATCTACACGCTTCTGTGAAGC 960  
DB |||||  
QY 961 CACCTTCCCCCACAATCTCAGTTCCCTTGGCCCTCAAAACTCACAAGGTTTTTCTTCAGAT 1020  
DB |||||  
QY 961 CACCTTCCCCCACAATCTCAGTTCCCTTGGCCCTCAAAACTCACAAGGTTTTTCTTCAGAT 1020  
DB |||||  
QY 1021 CAGCTCTCTGGAGGCTGGCAGGAGTGGGAAGGGGTGTGCTAGAGGGAGAAGCCTGCCT 1080  
DB |||||  
QY 1021 CAGCTCTCTGGAGGCTGGCAGGAGTGGGAAGGGGTGTGCTAGAGGGAGAAGCCTGCCT 1080  
DB |||||  
QY 1081 TGTGGTGGGACCTGATATACCCCTGAGCTCTCGGGAATGCTTTTCTGGGAGGAGCTG 1140  
DB |||||  
QY 1081 TGTGGTGGGACCTGATATACCCCTGAGCTCTCGGGAATGCTTTTCTGGGAGGAGCTG 1140  
DB |||||  
QY 1141 GAGAGCTCTCTAAACCTCTTCTCCCCAGAGACTAGATTGGCTTTTGTGATGTGTGGC 1200

DB |||||  
QY 1141 GAGAGCTCTCTAAACCTCTTCTCCCCAGAGACTAGATTGCCCTTGTGTTTGTGATGTGTGGC 1200  
DB |||||  
QY 1201 CTGAGAAATTGATCCATTTCCTGCTGTGTCTCTGGGGCGGCTCTCTTCCCATCT 1260  
DB |||||  
QY 1201 CTGAGAAATTGATCCATTTCCTGCTGTGTCTCTGGGGCGGCTCTCTTCCCATCT 1260  
DB |||||  
QY 1261 CCACCCCGCCCGCCAGAGCCATTGAGTGAGGTGCTTTTAGCCCTTTTGACTTAACATAA 1320  
DB |||||  
QY 1261 CCACCCCGCCCGCCAGAGCCATTGAGTGAGGTGCTTTTAGCCCTTTTGACTTAACATAA 1320  
DB |||||  
QY 1321 AATGACGCTGCTTGGGATAACGAGGCAAGACCTCTCTCCACCTGTGGGCTTGCCCAAG 1380  
DB |||||  
QY 1321 AATGACGCTGCTTGGGATAACGAGGCAAGACCTCTCTCCACCTGTGGGCTTGCCCAAG 1380  
DB |||||  
QY 1381 CCCCCACTCTGTGCTGAATGTTCTCTGAGGCTCTGGCTAGAGTCCAGGCCCAACCCAG 1440  
DB |||||  
QY 1381 CCCCCACTCTGTGCTGAATGTTCTCTGAGGCTCTGGCTAGAGTCCAGGCCCAACCCAG 1440  
DB |||||  
QY 1441 GAGGAGGAGCGGAGCTGGGAAAGTCCACCTGGAGAGCTGAGCGCTCTTGGGCTT 1500  
DB |||||  
QY 1441 GAGGAGGAGCGGAGCTGGGAAAGTCCACCTGGAGAGCTGAGCGCTCTTGGGCTT 1500  
DB |||||  
QY 1501 AGCACCAACCCAGATCCTTCTCCACCCCTCTCCCTGGCTCCATGTGACCATGACTGAGGG 1560  
DB |||||  
QY 1501 AGCACCAACCCAGATCCTTCTCCACCCCTCTCCCTGGCTCCATGTGACCATGACTGAGGG 1560  
DB |||||  
QY 1561 ACCAACTGGGCGCCACGCTAGGTGCGCCAGAGCTGTTAATGACTTCACTGCTTCC 1620  
DB |||||  
QY 1561 ACCAACTGGGCGCCACGCTAGGTGCGCCAGAGCTGTTAATGACTTCACTGCTTCC 1620  
DB |||||  
QY 1621 TGCAAGATCAGCTGTGCTGCTTTTGGCTTGGGCTGGCCACAGGGTCCAGGGACTCTG 1680  
DB |||||  
QY 1621 TGCAAGATCAGCTGTGCTGCTTTTGGCTTGGGCTGGCCACAGGGTCCAGGGACTCTG 1680  
DB |||||  
QY 1681 GCCTTAGCCCAAGAGTGAAGAGCTTACAGCCAGCTATGGAGCCCTTGGGCTTCC 1740  
DB |||||  
QY 1681 GCCTTAGCCCAAGAGTGAAGAGCTTACAGCCAGCTATGGAGCCCTTGGGCTTCC 1740  
DB |||||  
QY 1741 CT 1742  
DB |||||  
QY 1741 CT 1742  
DB |||||

RESULT 2  
BD013799  
LOCUS Modified cDNA of rat bcl-x gene and modified protein.  
DEFINITION  
ACCESSION BD013799  
VERSION BD013799.1 GI:22554128  
KEYWORDS JP 2001120281-A/1.  
SOURCE Rattus norvegicus  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
1 (bases 1 to 1742)  
Ota, S. and Abo, S.  
Modified cDNA of rat bcl-x gene and modified protein  
Patent: JP 2001120281-A 1 08-MAY-2001;  
JAPAN SCIENCE AND TECHNOLOGY CORP  
OS Rattus norvegicus (rat)  
PN JP 2001120281-A/1  
PD 08-MAY-2001  
PF 16-AUG-2000 JP 2000246999  
PI SHIGEO OTA, SADAMITSU ASO  
PC C12N15/09, C07K14/47, C12N1/15, C12N1/19, C12N1/21, C12N5/10// PC  
C12P21/02,  
PC (C12P21/02, C12R1:19), (C12P21/02, C12R1:91), C12N15/00, C12N5/00  
CC Modified cDNA of rat bcl-x gene and modified protein FH Key  
Location/Qualifiers  
(72). (773).  
FT CDS  
Location/Qualifiers

Qy	961	CACCTTCCCCCACAATCTCAGTTCCCTTGGCCTCAAAACTCAAGGTTTTTCTTCAGAT	1020
Db	961	CACCTTCCCCCACAATCTCAGTTCCCTTGGCCTCAAAACTCAAGGTTTTTCTTCAGAT	1020
Qy	1021	CAGCTCCTTGGAGGCTGGCAGGAGTGCGAAGGGTGCTAGAGGGAGAAGAGCGCTCGCT	1080
Db	1021	CAGCTCCTTGGAGGCTGGCAGGAGTGCGAAGGGTGCTAGAGGGAGAAGAGCGCTCGCT	1080
Qy	1081	TGTTGGTGGACCTCGATTACCCCTCAGCCTCTCGGGAATGCTTTTCTGCGAGGAGCTG	1140
Db	1081	TGTTGGTGGACCTCGATTACCCCTCAGCCTCTCGGGAATGCTTTTCTGCGAGGAGCTG	1140
Qy	1141	GAGAGCTCTCTAAACCTCTTCCCCAGAGACTAGATTGCCCTTGGTTTTGATGTGTGGC	1200
Db	1141	GAGAGCTCTCTAAACCTCTTCCCCAGAGACTAGATTGCCCTTGGTTTTGATGTGTGGC	1200
Qy	1201	CTCAGAATTCATCCATTTCCCATCTTGCTGTGTCCCTGGGCGGCTCCTCTCTCCCATCT	1260
Db	1201	CTCAGAATTCATCCATTTCCCATCTTGCTGTGTCCCTGGGCGGCTCCTCTCTCCCATCT	1260
Qy	1261	CCACCCCCCCCCCAGAGCCATTGAGTGAGGTGCTTTTAGCCCTTTTGACTAACTAAA	1320
Db	1261	CCACCCCCCCCCCAGAGCCATTGAGTGAGGTGCTTTTAGCCCTTTTGACTAACTAAA	1320
Qy	1321	AATGCAGGCTGCTTGGGATAACGAGGCAAGGACCTCTCCCACTGTGGCTTGGCCAAAG	1380
Db	1321	AATGCAGGCTGCTTGGGATAACGAGGCAAGGACCTCTCCCACTGTGGCTTGGCCAAAG	1380
Qy	1381	CCCCCACTCTGCTGTAATGTTCTCTGAGGCTCTGGCTAGAGTCCAGCCCAACCCAG	1440
Db	1381	CCCCCACTCTGCTGTAATGTTCTCTGAGGCTCTGGCTAGAGTCCAGCCCAACCCAG	1440
Qy	1441	GAGAGGAGCGAGCTGCGGAAAGTCACCTCGAGAGCTGAGCGGCTCTTCGGGCTT	1500
Db	1441	GAGAGGAGCGAGCTGCGGAAAGTCACCTCGAGAGCTGAGCGGCTCTTCGGGCTT	1500
Qy	1501	AGCACCAACCCAGATCTTCTCAACCCCTCCCTGGCTCCATGGTGAACCATGTAGGG	1560
Db	1501	AGCACCAACCCAGATCTTCTCAACCCCTCCCTGGCTCCATGGTGAACCATGTAGGG	1560
Qy	1561	ACCAACTGGGCCCAACGCTAGTGTCGCCAGAGCTGTTAATGACTTCAGCTGCCTCAGTCC	1620
Db	1561	ACCAACTGGGCCCAACGCTAGTGTCGCCAGAGCTGTTAATGACTTCAGCTGCCTCAGTCC	1620
Qy	1621	TGCAAGATCAGCTGTGGCATCTTTGCTTTGGGTGCTGGCCACAGGGTCCAGGGACTCTG	1680
Db	1621	TGCAAGATCAGCTGTGGCATCTTTGCTTTGGGTGCTGGCCACAGGGTCCAGGGACTCTG	1680
Qy	1681	GCCTTAGCCCAAGTGAGAGGAAGCTTACAGCGCAGCTATAGGAGCCCTTGGGGCTTCC	1740
Db	1681	GCCTTAGCCCAAGTGAGAGGAAGCTTACAGCGCAGCTATAGGAGCCCTTGGGGCTTCC	1740
Qy	1741	CT 1742	
Db	1741	CT 1742	
RESULT 3			
LOCUS	RNU72350	1742 bp	linear
DEFINITION	Rattus norvegicus Bcl-xalpha mRNA, complete cds.		
ACCESSION	U72350		
VERSION	U72350.1	GI:1622936	
KEYWORDS			
SOURCE	Rattus norvegicus (Norway rat)		
ORGANISM	Rattus norvegicus		
REFERENCE	1 (bases 1 to 1742)		
AUTHORS	Shiraiwa,N., Inohara,N., Okada,S., Yuzaki,M., Shoji,S. and Ohta,S.		
TITLE	An additional form of rat Bcl-x, Bcl-xbeta, generated by an		
	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;		
	Rattus.		
	ROD 22-OCT-1996		

unspliced RNA, promotes apoptosis in promyeloid cells  
J. Biol. Chem. 271 (22), 13258-13265 (1996)

JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

2 (bases 1 to 1742)  
Shiraiwa,N., Inohara,N., Okada,S., Yuzaki,M., Shoji,S. and Ohta,S.  
Direct Submission  
Submitted (24-SEP-1996) Division of Biochemistry, Institute of  
Gerontology, 1-396 Kosugi-cho, Nakahara-ku, Kawasaki City 211,  
Japan

## FEATURES

source Location/Qualifiers

1..1742

/organism="Rattus norvegicus"

/mol\_type="mRNA"

/db\_xref="taxon:10116"

72..773

/codon\_start=1

/product="Bcl-xalpha"

/protein\_id="AAB17353.1"

/db\_xref="GI:1622937"

/translation="MSQSNRELVDPLSYKLSQKYSQSPSDVEENRTEAPERTEPE  
RETFSALNGNSFWHLAUSPAVNGATGHSSSLDAREVIPMAVKQALREAGDEPELRVR  
RAFSDLTSQLHTPTAYQSPFQVWNLFRDGVNMRIVAFPSFGALCVESVDKMQ  
VLVSRIASWATYLNLDLEPWIQENGWDTPVDLYGNAAESRKGQERFNWPLTGM  
TVAGVLLGSLFSRK"

## ORIGIN

Query Match 100.0%; Score 1742; DB 10; Length 1742;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1742; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CACAGAGACACCCAGTCAGTGAGCAGGTGTTGGCAATGGACTGGTCAGGCCATCT 60  
Db |||||  
Qy 1 CACAGAGACACCCAGTCAGTGAGCAGGTGTTGGCAATGGACTGGTCAGGCCATCT 60  
Db |||||  
Qy 61 CTATTATATAAATGTCTCAGAGCAACCGGAGCTGGTGTGACTTCTCTCTCTACAAGC 120  
Db |||||  
Qy 61 CTATTATATAAATGTCTCAGAGCAACCGGAGCTGGTGTGACTTCTCTCTACAAGC 120  
Db |||||  
Qy 121 TCTCCAGAAAGGATACAGCTGGAGTCAGTTAGCGATGTGGAAGAACAGGACTGAAG 180  
Db |||||  
Qy 121 TCTCCAGAAAGGATACAGCTGGAGTCAGTTAGCGATGTGGAAGAACAGGACTGAAG 180  
Db |||||  
Qy 181 CCCAGAGAAACCTGAAACAGAAAGGAGACCCCGAGTGCCATCAATGGCAACCCATCCT 240  
Db |||||  
Qy 181 CCCAGAGAAACCTGAAACAGAAAGGAGACCCCGAGTGCCATCAATGGCAACCCATCCT 240  
Db |||||  
Qy 241 GGCACCTGGCGGATAGCCCGCGGTGAATGGAGCCACTGGCCACAGCAGCAGTTGGATG 300  
Db |||||  
Qy 241 GGCACCTGGCGGATAGCCCGCGGTGAATGGAGCCACTGGCCACAGCAGCAGTTGGATG 300  
Db |||||  
Qy 301 CGCGGAGGATATCCCATGGCAGCAGTGAAGAACAGCGCTGAGAGAGCTGGCGATGAGT 360  
Db |||||  
Qy 301 CGCGGAGGATATCCCATGGCAGCAGTGAAGAACAGCGCTGAGAGAGCTGGCGATGAGT 360  
Db |||||  
Qy 361 TTGAACCTGGGTACCGGAGAGCATTCAGTGATCTTAACATCCAGCTTCATATAACCCAG 420  
Db |||||  
Qy 361 TTGAACCTGGGTACCGGAGAGCATTCAGTGATCTTAACATCCAGCTTCATATAACCCAG 420  
Db |||||  
Qy 421 GGACAGCATATCAGAGCTTTGAAACAGGTAGTGAATGAATCTTTTCGGGATGGGGTAAACT 480  
Db |||||  
Qy 421 GGACAGCATATCAGAGCTTTGAAACAGGTAGTGAATGAATCTTTTCGGGATGGGGTAAACT 480  
Db |||||  
Qy 481 GGGGTGCGATTTGGCCCTCTCTCTTTGGCGGGCACTGTGCGTGGAAAGCGTAGACA 540  
Db |||||  
Qy 541 AGGAGATCCAGTATTCGTGATCGGATTCGAAATTCGATGGATGCCACCTACCTGAATGACC 600  
Db |||||  
Qy 541 AGGAGATCCAGTATTCGTGATCGGATTCGAAATTCGATGGATGCCACCTACCTGAATGACC 600  
Db |||||  
Qy 601 ACCTAGAGCTTTGGATCCAGAGAAACGGCGCTGGGACACTTTTGTGGATCTCTACGGGA 660  
Db |||||

Db 601 ACCTAGAGCTTTGGATCCAGAGAAACGGCGGCTGGGACACTTTTGTGGATCTCTACGGGA 660  
Qy |||||  
Db 661 ACAATGACAGCAGCGAGAGCGGAAAGGCCAGGAGCGTTTCAACCCGTGGTTCTGACGG 720  
Qy |||||  
Db 661 ACAATGACAGCAGCGAGAGCGGAAAGGCCAGGAGCGTTTCAACCCGTGGTTCTGACGG 720  
Qy |||||  
Db 721 GCATGACTGTGGTGGTGTAGTTCTGCTGGGCTCAGCTCTTTCAGTGGAAAGTGACGACACA 780  
Qy |||||  
Db 721 GCATGACTGTGGTGGTGTAGTTCTGCTGGGCTCAGCTCTTTCAGTGGAAAGTGACGACACA 780  
Qy |||||  
Db 781 CTGACCGTCCACTCACCTCTCACTCCCACTTGGCCCCCACCACTCTCTCTTCAGCC 840  
Qy |||||  
Db 781 CTGACCGTCCACTCACCTCTCACTCCCACTTGGCCCCCACCACTCTCTCTTCAGCC 840  
Qy |||||  
Db 841 ACCATTGCTACCAAGGAGAACCACTACATGCACTCAGCGCCCTTCCCTATTATAGGGTT 900  
Qy |||||  
Db 841 ACCATTGCTACCAAGGAGAACCACTACATGCACTCAGCGCCCTTCCCTATTATAGGGTT 900  
Qy |||||  
Db 901 GGGCTTAGACGGAGTCCCTTGCAGTTAGCTTTCTAGAAATCTTACCAAGCTTCTGTGAAAGC 960  
Qy |||||  
Db 901 GGGCTTAGACGGAGTCCCTTGCAGTTAGCTTTCTAGAAATCTTACCAAGCTTCTGTGAAAGC 960  
Qy |||||  
Db 961 CACCTTCCCCCACAATCTCAGTTCCCTTGGCCTCAAAACTCACAAGGTTTTTCTCAGAT 1020  
Qy |||||  
Db 961 CACCTTCCCCCACAATCTCAGTTCCCTTGGCCTCAAAACTCACAAGGTTTTTCTCAGAT 1020  
Qy |||||  
Db 1021 CAGCTCTTGGAGCTGGCAGGAGTGGGAAAGGGTGTCTAGAGGAGAGAGCTGCCT 1080  
Qy |||||  
Db 1021 CAGCTCTTGGAGCTGGCAGGAGTGGGAAAGGGTGTCTAGAGGAGAGAGCTGCCT 1080  
Qy |||||  
Db 1081 TGTGTGGGAGACCTGATTACCTCCTGAGCTCTCGGAAATGCTTTCTTGGCAGGAGCTG 1140  
Qy |||||  
Db 1081 TGTGTGGGAGACCTGATTACCTCCTGAGCTCTCGGAAATGCTTTCTTGGCAGGAGCTG 1140  
Qy |||||  
Db 1141 GAGAGCTCTCTAAACCTCTTCCCCCAGAGACTAGATTGCTTGGTGTGTGTGGC 1200  
Qy |||||  
Db 1141 GAGAGCTCTCTAAACCTCTTCCCCCAGAGACTAGATTGCTTGGTGTGTGTGGC 1200  
Qy |||||  
Db 1201 CTGAAATTGATTCATTTCCATTTCTGCTGTCTCTTGGGCGGCTCCTCTTCCATCT 1260  
Qy |||||  
Db 1201 CTGAAATTGATTCATTTCCATTTCTGCTGTCTCTTGGGCGGCTCCTCTTCCATCT 1260  
Qy |||||  
Db 1261 CCACCCCCCCCCAGAGCATTGAGTGAGGTGCTTTTAGCCCTTTTGTACTAATCTAAA 1320  
Qy |||||  
Db 1261 CCACCCCCCCCCAGAGCATTGAGTGAGGTGCTTTTAGCCCTTTTGTACTAATCTAAA 1320  
Qy |||||  
Db 1321 AATGACGGCTGTGGGATTAACGAGCAAGGAACTCTCTCCCACTGTGGCTGGCCAAAG 1380  
Qy |||||  
Db 1321 AATGACGGCTGTGGGATTAACGAGCAAGGAACTCTCTCCCACTGTGGCTGGCCAAAG 1380  
Qy |||||  
Db 1381 CCCCCACTCTGTGCTGAATGTTCTCTGAGGCTCTGGCTAGAGTCCAGCCCAACCCAG 1440  
Qy |||||  
Db 1381 CCCCCACTCTGTGCTGAATGTTCTCTGAGGCTCTGGCTAGAGTCCAGCCCAACCCAG 1440  
Qy |||||  
Db 1441 GAGGAGGAGCAGGAGCTCGGAAAGTCCACCTGGAGAGCCTGAGCGCTCTTGGCGGCTT 1500  
Qy |||||  
Db 1441 GAGGAGGAGCAGGAGCTCGGAAAGTCCACCTGGAGAGCCTGAGCGGCTCTTGGCGGCTT 1500  
Qy |||||  
Db 1501 AGCAACCCAGATCTTCTTCCACCCCTCCCTGGCTCCATGTGACCATGACTGAGGG 1560  
Qy |||||  
Db 1501 AGCAACCCAGATCTTCTTCCACCCCTCCCTGGCTCCATGTGACCATGACTGAGGG 1560  
Qy |||||  
Db 1561 ACCAACTGGGCCCAGCTAGGTGGCCAGAGCTGTTAATGACTTCAGCTGCTCACTTCC 1620  
Qy |||||  
Db 1561 ACCAACTGGGCCCAGCTAGGTGGCCAGAGCTGTTAATGACTTCAGCTGCTCACTTCC 1620  
Qy |||||  
Db 1621 TGCAAGATCAGCTGTGGCATCTTTTGGCTTGGGTGCTGGCCACAGGGTCCAGGACCTCTG 1680  
Qy |||||  
Db 1621 TGCAAGATCAGCTGTGGCATCTTTTGGCTTGGGTGCTGGCCACAGGGTCCAGGACCTCTG 1680  
Qy |||||  
Db 1681 GCCTTAGCCCAAGAGTGAAGGAAAGCTTACAGCCAGCTATGGAGCCCTCTGGGGCTTCC 1740  
Qy |||||  
Db 1681 GCCTTAGCCCAAGAGTGAAGGAAAGCTTACAGCCAGCTATGGAGCCCTCTGGGGCTTCC 1740

```
Oy 1741 CT 1742
Db 1741 CT 1742

RESULT 4
LOCUS RNCBLXLS 2232 bp DNA linear ROD 31-DEC-1994
DEFINITION R.norvegicus bcl-x gene.
ACCESSION X82537
VERSION X82537.1 GI:607176
KEYWORDS bcl-x gene.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE 1 (bases 1 to 2232)
AUTHORS Michaelidis,T.M.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2232)
AUTHORS Michaelidis,T.M.
TITLE Direct Submission
JOURNAL Submitted (04-NOV-1994) T.M. Michaelidis, Max-Planck-Inst. for
            Psychiatry, Dept of Neurochemistry, 82152 Martinsried, FRG
FEATURES             source
     source            /organism="Rattus norvegicus"
     source            /mol_type="genomic DNA"
     source            /db_xref="taxon:10116"
     source            /clone="X9, X31"
     source            /tissue type="brain"
     source            /clone_lib="hippocampal cDNA libr. and genomic DNA libr."
     source            /dev stage="embryonic"
     source            <1..19
     source            20..721
     source            /gene="bcl-x"
     source            /gene="bcl-x"
     source            /codon_start=1
     source            /protein_id="CAA57886.1"
     source            /db_xref="GI:607177"
     source            /db_xref="GOA:P53563"
     source            /db_xref="UniProt/Swiss-Prot:P53563"
     source            /translation="MSQSNQELVDFLSYKLSQKGYWSQPSDVEENRTEAPEPEPE
            RETPSALNGNSWLHADS PAVNGATGHSSSLDAREVIPMAAVKQALREAGDEFLRYR
            RAFSDLTSQLHITPTATAYQSFEQVWNLFRDGVNWRIVAFPSFGALCVESVDKMQ
            VLVSRIASWMTATYLDHLEPWIQENGWDTFVDLYGNNAASRSKQERFNRWFLTGM
            TVAGVLLGSLFSRK"
     source            join(20..394,584..721)
     source            /gene="bcl-x"
     source            /note="smaller form due to splicing"
     source            /codon_start=1
     source            /protein_id="CAA57887.1"
     source            /db_xref="GI:607178"
     source            /db_xref="GOA:P53563"
     source            /db_xref="UniProt/Swiss-Prot:P53563"
     source            /translation="MSQSNQELVDFLSYKLSQKGYWSQPSDVEENRTEAPEPEPE
            RETPSALNGNSWLHADS PAVNGATGHSSSLDAREVIPMAAVKQALREAGDEFLRYR
            RAFSDLTSQLHITPTATAYQSFEQVWNLFRDGVNWRIVAFPSFGALCVESVDKMQ
            VLVGSLFSRK"
     source            722..2232
     source            2184..2189
     source            polyA_signal

3'UTR
polyA_signal

ORIGIN
Query Match 95.4%; Score 1662.2; DB 10; Length 2232;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1686; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

Oy 53 GCCATCTCTATTATAAAATGCTCTCAGAGCAACCGGAGCTGGTGGTGTGACTTCTCTC 112
Db 1 GCGCATCTCTATTATAAAATGCTCTCAGAGCAACCGGAGCTGGTGGTGTGACTTCTCTC 60

113 CTCAAGCTCTCCAGAAAGGATACAGCTGAGTCACTTTAGCGATGTGGAAGAGAACAG 172
    |||
61 CTCAAGCTCTCCAGAAAGGATACAGCTGAGTCACTTTAGCGATGTGGAAGAGAACAG 120

173 GACTGAAGCCCAAGAAACTGAACAGAAAGGAGACCCCGAGTGCCTCAATGGCAA 232
    |||
121 GACTGAAGCCCAAGAAACTGAACAGAAAGGAGACCCCGAGTGCCTCAATGGCAA 180

233 CCCATCTCGGCACCTGGCGGATAGCCCGCGGTGAATGGAGCCACTGGCCACAGCAGCAG 292
    |||
181 CCCATCTCGGCACCTGGCGGATAGCCCGCGGTGAATGGAGCCACTGGCCACAGCAGCAG 240

293 TTGGATCGCGGAGAGTAATCCCATGCGCAGCAGTAGAAGCAAGCGCTGAGAGAGGCTGG 352
    |||
241 TCTGGATCGCGGAGAGTAATCCCATGCGCAGCAGTAGAAGCAAGCGCTGAGAGAGGCTGG 300

353 CGATGAGTTTGAACCTGCGGTACCGGAGAGCACTTCAATCATCTTAACATCCAGCTTCATAT 412
    |||
301 CGATGAGTTTGAACCTGCGGTACCGGAGAGCACTTCAATCATCTTAACATCCAGCTTCATAT 360

413 AACCCAGGGACAGCATATCAGAGCTTTGAACAGGTAGTGAATGAATCTTTTCGGGATGG 472
    |||
361 AACCCAGGGACAGCATATCAGAGCTTTGAACAGGTAGTGAATGAATCTTTTCGGGATGG 420

473 GGTAAACTGGGGTCGCATTTGGGCTTCTTCTCTTTGGGGGGGCACTGTGCGTGGAAAG 532
    |||
421 GGTAAACTGGGGTCGCATTTGGGCTTCTTCTCTTTGGGGGGGCACTGTGCGTGGAAAG 480

533 CGTAGACAGGAGATGAGGTATTTGTCAGTCGGATTCAGATTTGGATGGCCACCTACCT 592
    |||
481 CGTAGACAGGAGATGAGGTATTTGTCAGTCGGATTCAGATTTGGATGGCCACCTACCT 540

593 GAATGACCACTAGAGCTTGGATCCAGAGAAACGGGGCTGGGACACTTTTGTGGATCT 652
    |||
541 GAATGACCACTAGAGCTTGGATCCAGAGAAACGGGGCTGGGACACTTTTGTGGATCT 600

653 CTACGGGAAACAAATGCAGCAGCGGAGACCGGAAAGGCGCAGAGGCTTTCAAACCGTGGTT 712
    |||
601 CTACGGGAAACAAATGCAGCAGCGGAGACCGGAAAGGCGCAGAGGCTTTCAAACCGTGGTT 660

713 CCTACGGGATGACTGTGGCTGTGTAGTTCCTGCTGGGCTCACTCTTCAGTCGGAGTG 772
    |||
661 CCTACGGGATGACTGTGGCTGTGTAGTTCCTGCTGGGCTCACTCTTCAGTCGGAGTG 720

773 ACCAGACACTGACCGTCCACTCACTCTCACTCCACTTGCCTCCACCAACCACTCTCT 832
    |||
721 ACCAGACACTGACCGTCCACTCACTCTCACTCCACTTGCCTCCACCAACCACTCTCT 780

833 CTTCAGCCACCATTTGTACCAAGGAGAACCACTACATGCACTCAACCGCCCTTCCCTTATT 892
    |||
781 CTTCAGCCACCATTTGTACCAAGGAGAACCACTACATGCACTCAACCGCCCTTCCCTTATT 840

893 ATAGGTTGGGCTTAGACGAGTCCCTGCGAGTCTTCTAGATCTACACGCTTCT 952
    |||
841 ATAGGTTGGGCTTAGACGAGTCCCTGCGAGTCTTCTAGATCTACACGCTTCT 900

953 GTGAAGCCACTTTCCCCCACTCTCAGTTCCTTGGCTTCCCTCAAACTCAAAAGTTTTT 1012
    |||
901 GTGAAGCCACTTTCCCCCACTCTCAGTTCCTTGGCTTCCCTCAAACTCAAAAGTTTTT 960

1013 CCTCAGATCAGCTCTTTGGAGGCTGGCAGGAGTGGGAGGGGTGTGTAGAGGAGAGA 1072
    |||
961 CCTCAGATCAGCTCTTTGGAGGCTGGCAGGAGTGGGAGGGGTGTGTAGAGGAGAGA 1020

1073 GCCTGCTTGTGTGGGACCTGATTAACCTGAGCTCTCGGGAATGCTTTTCTGCA 1132
    |||
1021 GCCTGCTTGTGTGGGACCTGATTAACCTGAGCTCTCGGGAATGCTTTTCTGCA 1080

1133 GGGAGCTGGAGAGCTCTTAAACCTTTCCCCAGAGACTAGATTGCTTGGTTTGTATG 1192
    |||
1081 GGGAGCTGGAGAGCTCTTAAACCTTTCCCCAGAGACTAGATTGCTTGGTTTGTATG 1140
```

```
1193 TGTGTGGCTCAGAAATTGATCATTTCCATTCTGTGTGTCCTTGGGGCGGCTCCTCCT 1252
1141 TGTGTGGCTCAGAAATTGAT -CATTTCCATTCTGTGTGTCCTTGGGGCGGCTCCTCCT 1199
1253 TCCCATCTCCA -CCCCCCCCCCCCCCCCCAGAGCCATTGAGTGAGTGCTTTTAGCCCTTTTGA 1311
1200 TCCCATCTCCA -CCCCCCCCCCCCCCCCCAGAGCCATTGAGTGAGTGCTTTTAGCCCTTTTGA 1259
1312 CTAACATAAATGACGGCTGCTTGGGATAACGAGCAAGAGACTCTCTCCACACCTGTGGC 1371
1260 CTAACATAAATGACGGCTGCTTGGGATAACGAGCAAGAGACTCTCTCCACACCTGTGGC 1319
1372 CTGGCCAGAGCCCACTCTGTCTCTGAATGTTCTCTCAGAGCTCTGGCTAGAGTCCAGC 1431
1320 CTGGCCAGAGCCCACTCTGTCTCTGAATGTTCTCTCAGAGCTCTGGCTAGAGTCCAGC 1379
1432 CCCACCCAGGAGGAGGAGCGAGCTGCGGAAGTCCACCCTGCGAGAGCCCTGAGCGGCTC 1491
1380 CCCACCCAGGAGGAGGAGCGAGCTGCGGAAGTCCACCCTGCGAGAGCCCTGAGCGGCTC 1439
1492 TTGGGGCTTAGACACACCCAGATCTTCTCAGACCCCTCCCTGCTCCTCATGGTGACCAT 1551
1440 TTGGGGCTTAGACACACCCAGATCTTCTCAGACCCCTCCCTGCTCCTCATGGTGACCAT 1499
1552 GACTGAGGAGCAACACTGGGCCCCAGCTAGGTGCCCCAGAGCTGTTAATGACTTCAGCTGC 1611
1500 GACTGAGGAGCAACACTGGGCCCCAGCTAGGTGCCCCAGAGCTGTTAATGACTTCAGCTGC 1559
1612 CTCACTTCTGCAAGATCAGCTGTGGCATCTTTGCTTGGGTGCTGCCACAGAGGTCCA 1671
1560 CTCACTTCTGCAAGATCAGCTGTGGCATCTTTGCTTGGGTGCTGCCACAGAGGTCCA 1619
1672 GGGACTTGGCTTAGACCCAGAGTGAGAGGAAGCTTACAGCGCAGCTATGGGAGCCCTG 1731
1620 GGGACTTGGCTTAGACCCAGAGTGAGAGGAAGCTTACAGCGCAGCTATGGGAGCCCTG 1679
1732 GGGGCTTCCCT 1742
1680 GGGGCTTCCCT 1690

RESULT 5
AC147787/C
LOCUS AC147787 172803 bp DNA linear HTG 13-JAN-2004
DEFINITION Rattus norvegicus chromosome 1 clone RP32-62M8 map q32, WORKING
DRAFT SEQUENCES, 20 unordered pieces.
ACCESSION AC147787
VERSION AC147787.2 GI:40804772
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 172803)
AUTHORS Taenzler,S., Monti,J., Gloeckner,G., Goesele,C., Baumgart,C.,
Huebner,N. and Platzter,M.
Rat chromosome 1 genomic sequence
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 172803)
AUTHORS Lagemann,D. and Platzter,M.
Direct Submission
JOURNAL Submitted (10-JAN-2004) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena, Thuringia 07745, Germany
REFERENCE 3 (bases 1 to 172803)
AUTHORS Lagemann,D. and Platzter,M.
Direct Submission
JOURNAL Submitted (13-JAN-2004) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena, Thuringia 07745, Germany
COMMENT On Jan 13, 2004 this sequence version replaced gi:40789103.
----- Genome Center
Center: Institute of Molecular Biotechnology
Center code: IMB
```

FEATURES  
source

```
Web site: http://genome.imb-jena.de/
Contact: gscj-submit@genome.imb-jena.de
----- Project Information
Center project name: RA53
Center clone name: RP32-62M8
----- Summary Statistics
Sequencing vector: pUC18; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 167156 bases at least Q40
Consensus quality: 168803 bases at least Q30
Consensus quality: 169765 bases at least Q20
Quality coverage: 6.51x
-----
Sequence Quality Assessment:
This entry has been annotated with sequence quality
estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than
1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the
Genbank flat file format but are available as part
of this entry's ASN.1 file.
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 20 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 511: contig of 511 bp in length
1 512: gap of unknown length
1 513: gap of unknown length
1 514: gap of unknown length
1 515: gap of unknown length
1 516: gap of unknown length
1 517: gap of unknown length
1 518: gap of unknown length
1 519: gap of unknown length
1 520: gap of unknown length
1 521: gap of unknown length
1 522: gap of unknown length
1 523: gap of unknown length
1 524: gap of unknown length
1 525: gap of unknown length
1 526: gap of unknown length
1 527: gap of unknown length
1 528: gap of unknown length
1 529: gap of unknown length
1 530: gap of unknown length
1 531: gap of unknown length
1 532: gap of unknown length
1 533: gap of unknown length
1 534: gap of unknown length
1 535: gap of unknown length
1 536: gap of unknown length
1 537: gap of unknown length
1 538: gap of unknown length
1 539: gap of unknown length
1 540: gap of unknown length
1 541: gap of unknown length
1 542: gap of unknown length
1 543: gap of unknown length
1 544: gap of unknown length
1 545: gap of unknown length
1 546: gap of unknown length
1 547: gap of unknown length
1 548: gap of unknown length
1 549: gap of unknown length
1 550: gap of unknown length
1 551: gap of unknown length
1 552: gap of unknown length
1 553: gap of unknown length
1 554: gap of unknown length
1 555: gap of unknown length
1 556: gap of unknown length
1 557: gap of unknown length
1 558: gap of unknown length
1 559: gap of unknown length
1 560: gap of unknown length
1 561: gap of unknown length
1 562: gap of unknown length
1 563: gap of unknown length
1 564: gap of unknown length
1 565: gap of unknown length
1 566: gap of unknown length
1 567: gap of unknown length
1 568: gap of unknown length
1 569: gap of unknown length
1 570: gap of unknown length
1 571: gap of unknown length
1 572: gap of unknown length
1 573: gap of unknown length
1 574: gap of unknown length
1 575: gap of unknown length
1 576: gap of unknown length
1 577: gap of unknown length
1 578: gap of unknown length
1 579: gap of unknown length
1 580: gap of unknown length
1 581: gap of unknown length
1 582: gap of unknown length
1 583: gap of unknown length
1 584: gap of unknown length
1 585: gap of unknown length
1 586: gap of unknown length
1 587: gap of unknown length
1 588: gap of unknown length
1 589: gap of unknown length
1 590: gap of unknown length
1 591: gap of unknown length
1 592: gap of unknown length
1 593: gap of unknown length
1 594: gap of unknown length
1 595: gap of unknown length
1 596: gap of unknown length
1 597: gap of unknown length
1 598: gap of unknown length
1 599: gap of unknown length
1 600: gap of unknown length
1 601: gap of unknown length
1 602: gap of unknown length
1 603: gap of unknown length
1 604: gap of unknown length
1 605: gap of unknown length
1 606: gap of unknown length
1 607: gap of unknown length
1 608: gap of unknown length
1 609: gap of unknown length
1 610: gap of unknown length
1 611: gap of unknown length
1 612: gap of unknown length
1 613: gap of unknown length
1 614: gap of unknown length
1 615: gap of unknown length
1 616: gap of unknown length
1 617: gap of unknown length
1 618: gap of unknown length
1 619: gap of unknown length
1 620: gap of unknown length
1 621: gap of unknown length
1 622: gap of unknown length
1 623: gap of unknown length
1 624: gap of unknown length
1 625: gap of unknown length
1 626: gap of unknown length
1 627: gap of unknown length
1 628: gap of unknown length
1 629: gap of unknown length
1 630: gap of unknown length
1 631: gap of unknown length
1 632: gap of unknown length
1 633: gap of unknown length
1 634: gap of unknown length
1 635: gap of unknown length
1 636: gap of unknown length
1 637: gap of unknown length
1 638: gap of unknown length
1 639: gap of unknown length
1 640: gap of unknown length
1 641: gap of unknown length
1 642: gap of unknown length
1 643: gap of unknown length
1 644: gap of unknown length
1 645: gap of unknown length
1 646: gap of unknown length
1 647: gap of unknown length
1 648: gap of unknown length
1 649: gap of unknown length
1 650: gap of unknown length
1 651: gap of unknown length
1 652: gap of unknown length
1 653: gap of unknown length
1 654: gap of unknown length
1 655: gap of unknown length
1 656: gap of unknown length
1 657: gap of unknown length
1 658: gap of unknown length
1 659: gap of unknown length
1 660: gap of unknown length
1 661: gap of unknown length
1 662: gap of unknown length
1 663: gap of unknown length
1 664: gap of unknown length
1 665: gap of unknown length
1 666: gap of unknown length
1 667: gap of unknown length
1 668: gap of unknown length
1 669: gap of unknown length
1 670: gap of unknown length
1 671: gap of unknown length
1 672: gap of unknown length
1 673: gap of unknown length
1 674: gap of unknown length
1 675: gap of unknown length
1 676: gap of unknown length
1 677: gap of unknown length
1 678: gap of unknown length
1 679: gap of unknown length
1 680: gap of unknown length
1 681: gap of unknown length
1 682: gap of unknown length
1 683: gap of unknown length
1 684: gap of unknown length
1 685: gap of unknown length
1 686: gap of unknown length
1 687: gap of unknown length
1 688: gap of unknown length
1 689: gap of unknown length
1 690: gap of unknown length
1 691: gap of unknown length
1 692: gap of unknown length
1 693: gap of unknown length
1 694: gap of unknown length
1 695: gap of unknown length
1 696: gap of unknown length
1 697: gap of unknown length
1 698: gap of unknown length
1 699: gap of unknown length
1 700: gap of unknown length
1 701: gap of unknown length
1 702: gap of unknown length
1 703: gap of unknown length
1 704: gap of unknown length
1 705: gap of unknown length
1 706: gap of unknown length
1 707: gap of unknown length
1 708: gap of unknown length
1 709: gap of unknown length
1 710: gap of unknown length
1 711: gap of unknown length
1 712: gap of unknown length
1 713: gap of unknown length
1 714: gap of unknown length
1 715: gap of unknown length
1 716: gap of unknown length
1 717: gap of unknown length
1 718: gap of unknown length
1 719: gap of unknown length
1 720: gap of unknown length
1 721: gap of unknown length
1 722: gap of unknown length
1 723: gap of unknown length
1 724: gap of unknown length
1 725: gap of unknown length
1 726: gap of unknown length
1 727: gap of unknown length
1 728: gap of unknown length
1 729: gap of unknown length
1 730: gap of unknown length
1 731: gap of unknown length
1 732: gap of unknown length
1 733: gap of unknown length
1 734: gap of unknown length
1 735: gap of unknown length
1 736: gap of unknown length
1 737: gap of unknown length
1 738: gap of unknown length
1 739: gap of unknown length
1 740: gap of unknown length
1 741: gap of unknown length
1 742: gap of unknown length
1 743: gap of unknown length
1 744: gap of unknown length
1 745: gap of unknown length
1 746: gap of unknown length
1 747: gap of unknown length
1 748: gap of unknown length
1 749: gap of unknown length
1 750: gap of unknown length
1 751: gap of unknown length
1 752: gap of unknown length
1 753: gap of unknown length
1 754: gap of unknown length
1 755: gap of unknown length
1 756: gap of unknown length
1 757: gap of unknown length
1 758: gap of unknown length
1 759: gap of unknown length
1 760: gap of unknown length
1 761: gap of unknown length
1 762: gap of unknown length
1 763: gap of unknown length
1 764: gap of unknown length
1 765: gap of unknown length
1 766: gap of unknown length
1 767: gap of unknown length
1 768: gap of unknown length
1 769: gap of unknown length
1 770: gap of unknown length
1 771: gap of unknown length
1 772: gap of unknown length
1 773: gap of unknown length
1 774: gap of unknown length
1 775: gap of unknown length
1 776: gap of unknown length
1 777: gap of unknown length
1 778: gap of unknown length
1 779: gap of unknown length
1 780: gap of unknown length
1 781: gap of unknown length
1 782: gap of unknown length
1 783: gap of unknown length
1 784: gap of unknown length
1 785: gap of unknown length
1 786: gap of unknown length
1 787: gap of unknown length
1 788: gap of unknown length
1 789: gap of unknown length
1 790: gap of unknown length
1 791: gap of unknown length
1 792: gap of unknown length
1 793: gap of unknown length
1 794: gap of unknown length
1 795: gap of unknown length
1 796: gap of unknown length
1 797: gap of unknown length
1 798: gap of unknown length
1 799: gap of unknown length
1 800: gap of unknown length
1 801: gap of unknown length
1 802: gap of unknown length
1 803: gap of unknown length
1 804: gap of unknown length
1 805: gap of unknown length
1 806: gap of unknown length
1 807: gap of unknown length
1 808: gap of unknown length
1 809: gap of unknown length
1 810: gap of unknown length
1 811: gap of unknown length
1 812: gap of unknown length
1 813: gap of unknown length
1 814: gap of unknown length
1 815: gap of unknown length
1 816: gap of unknown length
1 817: gap of unknown length
1 818: gap of unknown length
1 819: gap of unknown length
1 820: gap of unknown length
1 821: gap of unknown length
1 822: gap of unknown length
1 823: gap of unknown length
1 824: gap of unknown length
1 825: gap of unknown length
1 826: gap of unknown length
1 827: gap of unknown length
1 828: gap of unknown length
1 829: gap of unknown length
1 830: gap of unknown length
1 831: gap of unknown length
1 832: gap of unknown length
1 833: gap of unknown length
1 834: gap of unknown length
1 835: gap of unknown length
1 836: gap of unknown length
1 837: gap of unknown length
1 838: gap of unknown length
1 839: gap of unknown length
1 840: gap of unknown length
1 841: gap of unknown length
1 842: gap of unknown length
1 843: gap of unknown length
1 844: gap of unknown length
1 845: gap of unknown length
1 846: gap of unknown length
1 847: gap of unknown length
1 848: gap of unknown length
1 849: gap of unknown length
1 850: gap of unknown length
1 851: gap of unknown length
1 852: gap of unknown length
1 853: gap of unknown length
1 854: gap of unknown length
1 855: gap of unknown length
1 856: gap of unknown length
1 857: gap of unknown length
1 858: gap of unknown length
1 859: gap of unknown length
1 860: gap of unknown length
1 861: gap of unknown length
1 862: gap of unknown length
1 863: gap of unknown length
1 864: gap of unknown length
1 865: gap of unknown length
1 866: gap of unknown length
1 867: gap of unknown length
1 868: gap of unknown length
1 869: gap of unknown length
1 870: gap of unknown length
1 871: gap of unknown length
1 872: gap of unknown length
1 873: gap of unknown length
1 874: gap of unknown length
1 875: gap of unknown length
1 876: gap of unknown length
1 877: gap of unknown length
1 878: gap of unknown length
1 879: gap of unknown length
1 880: gap of unknown length
1 881: gap of unknown length
1 882: gap of unknown length
1 883: gap of unknown length
1 884: gap of unknown length
1 885: gap of unknown length
1 886: gap of unknown length
1 887: gap of unknown length
1 888: gap of unknown length
1 889: gap of unknown length
1 890: gap of unknown length
1 891: gap of unknown length
1 892: gap of unknown length
1 893: gap of unknown length
1 894: gap of unknown length
1 895: gap of unknown length
1 896: gap of unknown length
1 897: gap of unknown length
1 898: gap of unknown length
1 899: gap of unknown length
1 900: gap of unknown length
1 901: gap of unknown length
1 902: gap of unknown length
1 903: gap of unknown length
1 904: gap of unknown length
1 905: gap of unknown length
1 906: gap of unknown length
1 907: gap of unknown length
1 908: gap of unknown length
1 909: gap of unknown length
1 910: gap of unknown length
1 911: gap of unknown length
1 912: gap of unknown length
1 913: gap of unknown length
1 914: gap of unknown length
1 915: gap of unknown length
1 916: gap of unknown length
1 917: gap of unknown length
1 918: gap of unknown length
1 919: gap of unknown length
1 920: gap of unknown length
1 921: gap of unknown length
1 922: gap of unknown length
1 923: gap of unknown length
1 924: gap of unknown length
1 925: gap of unknown length
1 926: gap of unknown length
1 927: gap of unknown length
1 928: gap of unknown length
1 929: gap of unknown length
1 930: gap of unknown length
1 931: gap of unknown length
1 932: gap of unknown length
1 933: gap of unknown length
1 934: gap of unknown length
1 935: gap of unknown length
1 936: gap of unknown length
1 937: gap of unknown length
1 938: gap of unknown length
1 939: gap of unknown length
1 940: gap of unknown length
1 941: gap of unknown length
1 942: gap of unknown length
1 943: gap of unknown length
1 944: gap of unknown length
1 945: gap of unknown length
1 946: gap of unknown length
1 947: gap of unknown length
1 948: gap of unknown length
1 949: gap of unknown length
1 950: gap of unknown length
1 951: gap of unknown length
1 952: gap of unknown length
1 953: gap of unknown length
1 954: gap of unknown length
1 955: gap of unknown length
1 956: gap of unknown length
1 957: gap of unknown length
1 958: gap of unknown length
1 959: gap of unknown length
1 960: gap of unknown length
1 961: gap of unknown length
1 962: gap of unknown length
1 963: gap of unknown length
1 964: gap of unknown length
1 965: gap of unknown length
1 966: gap of unknown length
1 967: gap of unknown length
1 968: gap of unknown length
1 969: gap of unknown length
1 970: gap of unknown length
1 971: gap of unknown length
1 972: gap of unknown length
1 973: gap of unknown length
1 974: gap of unknown length
1 975: gap of unknown length
1 976: gap of unknown length
1 977: gap of unknown length
1 978: gap of unknown length
1 979: gap of unknown length
1 980: gap of unknown length
1 981: gap of unknown length
1 982: gap of unknown length
1 983: gap of unknown length
1 984: gap of unknown length
1 985: gap of unknown length
1 986: gap of unknown length
1 987: gap of unknown length
1 988: gap of unknown length
1 989: gap of unknown length
1 990: gap of unknown length
1 991: gap of unknown length
1 992: gap of unknown length
1 993: gap of unknown length
1 994: gap of unknown length
1 995: gap of unknown length
1 996: gap of unknown length
1 997: gap of unknown length
1 998: gap of unknown length
1 999: gap of unknown length
2000: gap of unknown length
```











Db	212	CCCCAGAGAAACTGAAGCAGAGAGGAGACCCCCAGTGCATCAATGCGCAACCCCATCTCT	271
Qy	241	GGCACCTGGCGGATAGCCCGCGGTGAATGGAGCCACTGGCCACAGCAGCAGTTTGGATG	300
Db	272	GGCACCTGGCGGATAGCCCGCGGTGAATGGAGCCACTGGCCACAGCAGCAGTTTGGATG	331
Qy	301	CGCGGAGGTAATCCCATCGCAGCAGTGAAGCAAGCGCTGAGAGAGGCTGGCGATGAGT	360
Db	332	CGCGGAGGTAATCCCATGGCAGCAGTGAAGCAAGCGCTGAGAGAGGCGAGCGATGAGT	391
Qy	361	TTGAACTGCGGTACCGGAGAGCATTCAGTGCATCTAACTCCAGCTTCATATAACCCACG	420
Db	392	TTGAACTGCGGTACCGGAGAGCGTTCACTGATCTTAACATCCAGCTTCACATAACCCACG	451
Qy	421	GGACAGCATATCAGAGCTTTTGAAACAGGTAGTGAATGAATCTTTTCGGGATGGGGTAAACT	480
Db	452	GGACCGGTATCAGAGCTTTTGACGAGTAGTGAATGAATCTTTTCGGGATGGAGTAAACT	511
Qy	481	GGGTCGCAATGTGGCTTCTTCTCTTTTCGCGGGCACTGTGCTGGAAAGCGTAGACA	540
Db	512	GGGTCGCACTGTGGCTTTTCTCTCTTTTCGCGGGCACTGTGCTGGAAAGCGTAGACA	571
Qy	541	AGGAGATGACAGGTATTTGGTAGTCGATTCGAATGGATGGCCACTTACTGAATGACG	600
Db	572	AGGAGATGACAGGTATTTGGTAGTCGATTCGAATGGATGGCCACTTACTGAATGACG	631
Qy	601	ACCTAGAGCCTTGGATCCAGAGAAACGGCGCTGGGACACTTTTGTGGATCTCTACGGGA	660
Db	632	ACCTAGAGCCTTGGATCCAGAGAAACGGCGCTGGGACACTTTTGTGGATCTCTACGGGA	691
Qy	661	ACAATGACAGACCGCAGAGCGGAAAGCGCAGGAGCGTTTCAACCGCTGGTCTCTGACGG	720
Db	692	ACAATGACAGACCGCAGAGCGGAAAGCGCAGGAGCGTTTCAACCGCTGGTCTCTGACGG	751
Qy	721	GCATGACTGTGGTGTTAGTTCTGTGGGCTCACTCTTTCAGTCGGAAGTAGCAGACA	780
Db	752	GCATGACTGTGGTGTTAGTTCTGTGGGCTCACTCTTTCAGTCGGAAGTAGCAGACA	811
Qy	781	CTGACCGTCCACTCACCTCTCACCTCCACTTGGCCCCCACCACCACTCTCTCTTCAGCC	840
Db	812	CTGACCGTCCACTCACCTCTCACCTCCACTTGGCCCCCACCACCACTCTCTCTTCAGCC	871
Qy	841	ACCAATTGCTACCGAGAAACCATACATGCAACTCACGCCCTTCCCTATTATAGGGTT	900
Db	872	ACCAATTGCTACCGAGAAACCATACATGCAACTCACGCCCTTCCCTATTACAGGGTT	931
Qy	901	GGGCTTAGACGGAGTCCCTTGAGTTAGCTTTCTAGAACTACACAGCTTCTGTGAAAGC	960
Db	932	GGGCTTAGACGGAGTCCCTTGCAATTAGCTTTCTTAGGATCTACACACTTCTGTGAAAGC	991
Qy	961	CACCTTCCCCCACAATCTCAGTTCCCTTGGCTCAAAACTCACAAGGTTTTTCTCAGAT	1020
Db	992	CACCTTCCCCCACAATCTCAGTTCTCTTGGCTCAAAACTCACAAGGTTTTTCTCAGAT	1051
Qy	1021	CAGCTCTTGGAGGCTGGCAGGAGTGGGAAGGGGTGTGCTAGAGGGAGAGAGCGCTGCCT	1080
Db	1052	CAGCTCTTGGAGGCTGGCAGGAACAAGAGGGGTGTG---TAGGGAGAGAGCGCTGTCT	1108
Qy	1081	TGTTGGTGGGACCTCGATTACCCCTGAGCC--TCTCGGGAATGCTTTTCTGGCAGGAGCT	1139
Db	1109	TGTTGGTGGGACCTTAATTACCCCTGAGCTTCTCGGGAATGCTTTTCCAGCAGGAGCT	1168
Qy	1140	GGAGAGCTCTAAACCTTCTTCCCCCAGACACTAGATTGCTTGGTTTTGATGTGTGTG	1199
Db	1169	-----CTTCCCTTGAAGGAGACTAGCTTGGCTTGGTTTTGATGTGTGTG	1213
Qy	1200	CCTCAGAAATGATCAATTTCCCATTTCTGTGTGTGTCCTCGGGGGCTCTCTTCCCATC	1259
Db	1214	CCTCAGAGGTGATC--ATTTCCTATCCGTGTGTCTCCCTGGG-----TCTCCCTCTC	1263
Qy	1260	TCCACCCCCCCCCCAGAGCCATTCAAGTAGGTGCTTTTAGCCCTTTTACTAACTAA	1319
Db	1264	CCCATCTCCACCCCCCAAGAGCCATTCAAGTAGGTGCTTTTAGCCCTTTTACTAATTA	1323

Qy	1320	AAATGCAGGCTGCTTGGGATAACAGGCCAAG-----ACCTCTCCCCACCTGTGGCCT	1373
Db	1324	AAATTTCAGGCTGCTTGGGATAATAGAGACAAGCACCAACTCTCTCCCACTGTGGCCT	1383
Qy	1374	GGCC-ANGCCCCCACTCTCGTCTGAATGTTCTCTGAGGCCCTTGCTAGATCCAGCC	1432
Db	1384	GGCAAACCCCACCTTGTGGTCTGAATGTTCTCTGAGGCCCTTGCTAGAGTCTACGC	1443
Qy	1433	CCACCAGGAGGAGGACGGAG	1454
Db	1444	TCACCAGGAGGAGGCTTAG	1465
RESULT 8			
MMU51278			
LOCUS	MMU51278	1466 bp mRNA linear ROD 21-NOV-1997	
DEFINITION	Mus musculus thymus Bcl-xL mRNA, complete cds.		
ACCESSION	U51278		
VERSION	U51278.1	GI:2636673	
KEYWORDS			
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;		
AUTHORS	Yang,X.F., Weber,G.P. and Cantor,H. 1 (bases 1 to 1466)		
TITLE	A novel Bcl-x isoform connected to the T cell receptor regulates apoptosis in T cells		
JOURNAL	Immunity 7 (5), 629-639 (1997)		
MEDLINE	98051053		
PUBMED	9390687		
REFERENCE	2 (bases 1 to 1466) Yang,X.-F. and Cantor,H. Direct Submission		
AUTHORS			
TITLE	Submitted (13-MAR-1996) Xiao-Fang Yang, Immunopathology, Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115,		
JOURNAL	USA		
FEATURES			
source	Location/Qualifiers		
	1..1466		
	/organism="Mus musculus"		
	/mol_type="mRNA"		
	/strain="B6/CBA"		
	/db_xref="taxon:10090"		
	/tissue_type="thymus"		
	103..804		
	/note="Bcl-x long form"		
	/codon_start=1		
	/product="Bcl-xL"		
	/protein_id="AAC53459.1"		
	/db_xref="GI:2636674"		
	/translation="MQSNRELVDPLSYKLQKYSWSQFSVEENRTEAPEETAE RETFSDALNPSWHLADSPAVNGATHSSLDAREVIPMAVKQALREAGDFELRYR RAFDSTLSQHITPGYAYQFEQVNLELPFRDWGNVRIVAFPFSGALCVESVDKEMO VLVSRIASWMATYLNDHLEPWIQENGWDFTVDLYGNMAAESRKKEGFNRFLTWGM TVAGVVLLGSLSRSK"		
CDS			
ORIGIN			
Query Match	67.0%; Score 1168; DB 10; Length 1466;		
Best Local Similarity	91.0%; Pred. No. 0;		
Matches 1331; Conservative	0; Mismatches 95; Indels 36; Gaps 7;		
Qy	1	CACAGAGCAGCCAGTGAGTGACAGGTCTTTTGGCAATGGACTGGTGTAGCCCATCT	60
Db	32	CACAGAGCAGCCAGTAAGTGACAGGTGTTTTGGCAATGGACTGGTGTAGCCCATCT	91
Qy	61	CTATTATAAAAATGCTCTCAGAGCAACCGGAGCTGGTGTTGACTTTCTCTCTACAAGC	120
Db	92	CTATTATAAAAATGCTCTCAGAGCAACCGGAGCTGGTGTTGACTTTCTCTCTACAAGC	151
Qy	121	TCTCCCAGAAAGGATACAGCTGGAGTCAGTTTAGCGATGTGCAAGAGAACHAGACTGAG	180
Db	152	TTTCCCAGAAAGGATACAGCTGGAGTCAGTTTAGTGATGTGCAAGAGAATAGGACTGAGG	211





```

Db 215598 CGAGAGCTGAGCGCTCTTGGCGCTTAGCACCACCCAGATCTTCTCCACCCCTCCCC 215657
Qy 1534 TGGCTCCATGTTGACCATGACTGAGGGACCAACTGGGCCACCGTAGTGCGCCAGAGCT 1593
Db 215658 TGGCTCCATGTTGACCATGACTGAGGGACCAACTGGGCCACCGTAGTGCGCCAGAGCT 215717
Qy 1594 GTTAATGACTTCAGCTGCTCCTCCTCTGCAAGATGAGCTGTGGCACTTTGCTTGGG 1653
Db 215718 GTTAATGACTTCAGCTGCTCCTCCTCTGCAAGATGAGCTGTGGCACTTTGCTTGGG 215777
Qy 1654 TGCTGGCCACAGGCTCAGGAGCTGCGCTTAGCCAGAGTGCAGAGGAGCTTACAGC 1713
Db 215778 TGCTGGCCACAGGCTCAGGAGCTGCGCTTAGCCAGAGTGCAGAGGAGCTTACAGC 215837
Qy 1714 GCAGCTATGGAGCCCTGGGGGCTTCCCT 1742
Db 215938 GCAGCTATGGAGCCCTGGGGGCTTCCCT 215966

RESULT 10
AC106107 249872 bp DNA linear HTG 10-MAY-2003
LOCUS Rattus norvegicus clone CH230-137H11, *** SEQUENCING IN PROGRESS
DEFINITION *** 8 unordered pieces.
ACCESSION AC106107
VERSION AC106107.5 GI:30521597
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 249872)
Muzny,D., Marie., Metzker,M., Lee., Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,Y., Aoyagi,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Cesar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyte,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregregis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpasch,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensuhewa,L., Loulseged,H., Lozano,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindratne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapus,P., Martin,K., Martin,R., Martinez,E.,
Mawhinney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nair,L.,
Nwankweliem,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfamkoch,C.,
Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.,
Puzo,M., Quiloz,J., Rachin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savary,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,

```

```

Sneed,A., Sodergren,E., Song,X.-Z., Sozelle,R., Sosa,J.,
Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Uemari,K.,
Valas,R., Vang,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,F., Zhang,J., Zhou,X., Zhou,S., Zhao,S., Dunn,D., von
Niederhauser,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.
Direct Submission
Unpublished
2 (bases 1 to 249872)
Worley,K.C.
Direct Submission
Submitted (12-JAN-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 249872)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (10-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On May 10, 2003 this sequence version replaced gi:24954343.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GHKL
Center clone name: CH230-137H11
----- Summary Statistics
Assembly program: Atlas 3.0
Consensus quality: 232472 bases at least Q40
Consensus quality: 235092 bases at least Q30
Consensus quality: 236668 bases at least Q20
Estimated insert size: 242767; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 238024: contig of 238024 bp in length
* 238025 238124: gap of unknown length
* 238125 239508: contig of 1382 bp in length
* 239507 239606: gap of unknown length
* 239607 240748: contig of 1142 bp in length
* 240749 240848: gap of unknown length
* 240849 242686: contig of 1838 bp in length
* 242687 242786: gap of unknown length
* 242787 244646: contig of 1860 bp in length
* 244647 244746: gap of unknown length

```

```
* 244747 245903: contig of 1157 bp in length
* 245904 246003: gap of unknown length
* 246004 247803: contig of 1800 bp in length
* 247804 247903: gap of unknown length
* 247904 249872: contig of 1969 bp in length.
FEATURES
    source      location/Qualifiers
        1..249872
            /organism="Rattus norvegicus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10116"
            /clones="CH230-137H11"
        1..1543
            /note="wgs contig"
        6888..8959
            /note="wgs contig"
        232899..233997
            /note="wgs contig"
        236068..238024
            /note="wgs contig"
        /note="wgs_contig"

misc_feature
misc_feature
misc_feature
misc_feature

ORIGIN
Query Match      62.3%; Score 1085; DB 2; Length 249872;
Best Local Similarity 99.8%; Pred. No. 1e-303;
Matches 1107; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

Qy 635 GGACACTTTTCTGGATCTCTACGGGACAATGACGACCGACGAGCCGGAAGGCCAGGA 694
Db 68124 GGACACTTTTCTGGATCTCTACGGGACAATGACGACCGACGAGCCGGAAGGCCAGGA 68183

Qy 695 GCGTTTCAACCGCTGGTTCTCTGACGGGCATGACGTGGCTGGTGTAGTTCTGTGGGCTC 754
Db 68184 GCGTTTCAACCGCTGGTTCTCTGACGGGCATGACGTGGCTGGTGTAGTTCTGTGGGCTC 68243

Qy 755 ACTCTTCTAGTGGAGTGACAGACACTGACGCGTCTCACTCACTCTCACTCCCACTTG 814
Db 68244 ACTCTTCTAGTGGAGTGACAGACACTGACGCGTCTCACTCACTCTCACTCCCACTTG 68303

Qy 815 CCCCCACACAACTCTCTTTCAGCCACCATTTGTTACGAGGAGAACCACTACATGCAACT 874
Db 68304 CCCCCACACAACTCTCTTTCAGCCACCATTTGTTACGAGGAGAACCACTACATGCAACT 68363

Qy 875 CACGCCCCCTCCCTATTATAGGGTTGGCCCTAGACGGAGTCCCTGTGAGTTAGTTTCT 934
Db 68364 CACGCCCCCTCCCTATTATAGGGTTGGCCCTAGACGGAGTCCCTGTGAGTTAGTTTCT 68423

Qy 935 AGAATCTACACGCTTCTGTGAAGGACACTTCCGCCCAATCTCAGTTCCCTTGGGCTC 994
Db 68424 AGAATCTACACGCTTCTGTGAAGGACACTTCCGCCCAATCTCAGTTCCCTTGGGCTC 68483

Qy 995 AAAACTCACAGGTTTTTCTCAGATCAGCTCCTTGGAGGCTGGCAGAGTGGGAGGG 1054
Db 68484 AAAACTCACAGGTTTTTCTCAGATCAGCTCCTTGGAGGCTGGCAGAGTGGGAGGG 68543

Qy 1055 TGTGCTAGAGGAGAGAGCGCTGCTTGTGGTGAGCCCTGATTTACCCCTGAGCCTCTC 1114
Db 68544 TGTGCTAGAGGAGAGAGCGCTGCTTGTGGTGAGCCCTGATTTACCCCTGAGCCTCTC 68603

Qy 1115 GGGAAATGCTTTTCTGGCAGGAGCTGTGAGAGCTCTCTAAACCTCTTCCCCCAGAGACTAG 1174
Db 68604 GGGAAATGCTTTTCTGGCAGGAGCTGTGAGAGCTCTCTAAACCTCTTCCCCCAGAGACTAG 68663

Qy 1175 ATTGCTTGTGTTTGTAGTGTGGCCCTCAGAAATGATCCATTTCCCATTTCTGCTGTGTC 1234
Db 68664 ATTGCTTGTGTTTGTAGTGTGGCCCTCAGAAATGAT-CAPTTTCCCATTTCTGCTGTGTC 68722

Qy 1235 CCTGGGGCGGCTCTCTCTTCCCATCTCCA-CCCCCCCCCCCCCAGAGCAATTGAGTGAGG 1293
Db 68723 CCTGGGGCGGCTCTCTCTTCCCATCTCCA-CCCCCCCCCCCCCAGAGCAATTGAGTGAGG 68782

Qy 1294 TGCTTTTAGCCCTTTTGTACTAACTAAATGAGAGGCTGCTTGGGATACAGAGGCAAGGAC 1353
Db 68783 TGCTTTTAGCCCTTTTGTACTAACTAAATGAGAGGCTGCTTGGGATACAGAGGCAAGGAC 68842
```

Db 536 GGCACCTGGCAGACAGCCCGCGGTGATGGAGCCACTGGCCACAGCAGCAGTTGGATG 595  
Qy CGCGGAGTAAATCCCATGTCAGCAGTGAAGAGCGCTCAGAGAGGCTGGCGATGATG 360  
Db CGCGGAGTAAATCCCATGTCAGCAGTGAAGAGCGCTCAGAGAGGCTGGCGATGATG 655  
Qy TTGAATCGCGGTACCGGCGGCAATTCAGTACCTGACATCCAGCTCCACATCACCCCGAG 715  
Db TTGAATCGCGGTACCGGCGGCAATTCAGTACCTGACATCCAGCTCCACATCACCCCGAG 715  
Qy GGAAGCATATCAGAGCTTTGAACAGGTAGTGAATGAATCTTTCCGGATGGGTAACT 480  
Db GGAAGCATATCAGAGCTTTGAACAGGTAGTGAATGAATCTTTCCGGATGGGTAACT 775  
Qy GGGGTGCGATTTGGCGCTTTCTCTCTTTGGCGGCACTGCTGCTGGAAGCGTAGACA 540  
Db GGGGTGCGATTTGGCGCTTTCTCTCTTTGGCGGCACTGCTGCTGGAAGCGTAGACA 835  
Qy AGGAGATGCAAGTATTTGGTGTAGTCGGAATTCGCAAGTTGGATGGCCACCTTACCTGAATGACC 600  
Db AGGAGATGCAAGTATTTGGTGTAGTCGGAATTCGCAAGTTGGATGGCCACCTTACCTGAATGACC 895  
Qy ACCTAGAGCTTTGATTCAGAGAAACGGCGGCTGGGACACTTTTGTGATCTCTACCGGA 660  
Db ACCTAGAGCTTTGATTCAGAGAAACGGCGGCTGGGACACTTTTGTGATCTCTACCGGA 955  
Qy ACAATGCAAGCAGCGAGACCGGAAGCCAGGAGCGTTTCAACCGCTGTTCTTGACGG 720  
Db ACAATGCAAGCAGCGAGACCGGAAGCCAGGAGCGTTTCAACCGCTGTTCTTGACGG 1015  
Qy GCATGACTGTGGCTGTAGTTCTGCTGGGCTCACTCTTCAGTCGGAAGTGACACAGACA 780  
Db GCATGACTGTGGCTGTAGTTCTGCTGGGCTCACTCTTCAGTCGGAAGTGACACAGACA 1075  
Qy CTGACCGTCCATCCTACCTTCCACCTTCCACCTTGGCCCCCACAACCTCTCTTCAGCC 840  
Db CTGACCGTCCATCCTACCTTCCACCTTGGCCCCCACAACCTCTCTTCAGCC 1135  
Qy ACCATTGCTACCGAGAACCACTACATGCAACTACAGCCCTTCCCTTATTTAGGGTT 900  
Db ACCATTGCTACCGAGAACCACTACATGCAACTACAGCCCTTCCCTTATTTAGGGTT 1195  
Qy GGGCTAGACGAGTCCCTGCTGAGTTAGTTCTTGAATCTACACAGCTTCTGTGAAGC 960  
Db GGGCTAGACGAGTCCCTGCTGAGTTAGTTCTTGAATCTACACAGCTTCTGTGAAGC 1253  
Qy CACCTTCCCCCACAATCTCAGTTCCCTTGGCTCMAAACTCACAAGTCTTTCCTCAGAT 1020  
Db CACCTTCCCCCACAATCTCAGTTCCCTTGGCTCMAAACTCACAAGTCTTTCCTCAGAT 1305  
Qy CAGCTCTTGGAGGCTGCGAGAGTGGGAAGGGTGTCTAGAGG--AGAAGAGCTTGC 1078  
Db CAGCTCTTGGAGGCTGCGAGAGTGGGAAGGGTGTCTAGAGG--AGAAGAGCTTGC 1365  
Qy CTGTCCTAAGAGGCTGCGAGGTTAGGAAGGGTTTGTGGCTGGGGCAGGAGGCGCTTAC 1138  
Db CTGTCCTAAGAGGCTGCGAGGTTAGGAAGGGTTTGTGGCTGGGGCAGGAGGCGCTTAC 1423  
Qy CTGTTGTGTGGAGCCCTGATTAACCTCTGAGCTCTCGGGAATGCTTTTCTGCGAGGAGC 1138  
Db CTGTTGTGTGGAGCCCTGATTAACCTCTGAGCTCTCGGGAATGCTTTTCTGCGAGGAGC 1423  
Qy TGGAGAGCTCTCTAAACCTCTTCCCTC---CAGAGACTAGATGCTTGGTTTGGATGTG 1194  
Db TGGAGAGCTCTCTAAACCTCTTCCCTC---CAGAGAGGAGACTAGATGCTTGGTTTGGATGTG 1483  
Qy TGTGGCTCAGAAATGATGATTTCCCAATTTCTGCTGCTCCCTGGGGGGCTCTCTCTTC 1254  
Db TGTGGCTCAGAAATGATGATTTCCCAATTTCTGCTGCTCCCTGGGGGGCTCTCTCTTC 1543  
Qy CCATCTCCACCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 1314  
Db CC--TTCCATCCCTTACCCCTTAAGAGCCATTTAGGGGCGCCTTT-----TG 1587  
Qy ACTAAATGAGGCTGCTGGGATTAAGGAGGAGGAGCT-----CCTCCCGAGCTGT 1368

Db 1588 ACTAGGATTCAGGCTGCTGGGATAAAGATGCAAGGACCAAGGACTCCTCCTCCTCCTCT 1647  
Qy GGCCTGGCCA-AGCCCCACTCCTGCTGTAATGTTCTCTGAGGCTCTGGCTAGAGTC 1427  
Db GGCCTGGCTAGCTCTCACTCCAGTCCAAATGTTCTCCAGAAAGCTCTGGCTAGAGC 1707  
Qy CAGCCCCACCCAGAG--GAGGACCGAGCTGGGAAAGTCCACCTTCGAGAGCCTGA 1484  
Db CAGCCCCACCCAGAGGAGGGGCTATAGCTACAGGAAGCACCCTCATGSCCAAAGCTAGG 1767  
Qy GCGCTCTTGGGCTTAGCACCCAGATCCCTTCTCCACCCCTCCCTCGCTCCATGG 1544  
Db GTGGCCCTTGAGTTACAGACCCCTAGTCCCTTCCCTCGCTCCCTC-----CC 1818  
Qy TGACCATGCTAGGAGCAACTGGGCGCCAGCTAGGTGCCCCCAGAGCTGTAATGACTT 1604  
Db ATGACCATGCTAGGAGCAACTGGGCGCCAGACAGATGCCCCAGAGCTGTTATGGCT 1878  
Qy CAGTGGCTCACTTCTGCAAGATCAGCTGTGGATCTTTGGCTTGGGTGCTGGCCACA 1664  
Db CAGTGGCTCACTTCTGCAAGATCAGCTGTGGATCTTTGGCTTGGGTGCTGGCCACA 1935  
Qy GGGTCCAGGAGCTTGGCTTAGCCAGAGTGGAGGAGCTTACAGCTTACAGCTATGG 1723  
Db CATGGTGGTTCAGGAGCTCAGCCCTGAGTGAAGGAGCTATCAGAAACAGCTATGG 1995  
Qy GAGCCCTGGGCGCTTCCCT 1742  
Db GAGCCCTGGGCGCTTCCCT 2014

BC019307 2575 bp mRNA linear PRI 29-JUN-2004  
LOCUS Homo sapiens BC12-like 1, transcript variant 1, mRNA (cdna clone  
DEFINITION MGC:4204 IMAGE:2823498), complete cds.  
ACCESSION BC019307  
VERSION MGC.  
KEYWORDS BC019307.1 GI:17939633  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

RESULT 12  
BC019307  
LOCUS Homo sapiens BC12-like 1, transcript variant 1, mRNA (cdna clone  
DEFINITION MGC:4204 IMAGE:2823498), complete cds.  
ACCESSION BC019307  
VERSION MGC.  
KEYWORDS BC019307.1 GI:17939633  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 2575)  
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,  
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,  
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,  
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heide, F.,  
Diachenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,  
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,  
Scheetz, T.E., Brownstein, M.J., Uedin, T.B., Towbin, T.,  
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,  
Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,  
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,  
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,  
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,  
Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,  
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,  
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,  
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,  
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smaluk, D.E.,  
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.  
Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
12477932  
2 (bases 1 to 2575)  
Strausberg, R.  
Direct Submission  
Submitted (13-DEC-2001) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA







```
QY 1315 ACTAAATGAGGCTGCTTGGGATACAGAGGCAAGGACCT-----CCTCCCACTCTGT 1368
Db 1588 ACTAGGATTCAGGCTGCTTGGGATAAAGATGCAAGGACAGGACTCCCTCTCACCTCT 1647
QY 1369 GGCCTGGGCA-AGCCCCACCTCTGCTGCTGAATGTTCTCTGAGGCTCTCTGGTAGAGTC 1427
Db 1648 GGAATGGCTAGAGTCTCACTCCCACTGCTCAATGCTCTCCAGAGGCTCTCTGGTAGAGGC 1707
QY 1428 CAGCCCCACCCAGGAG---GAGGGACGAGCTGCGGAAAGTCCACCTTCGAGAGCCTTGA 1484
Db 1708 CAGCCCCACCCAGGAGGAGGGGCTATAGCTACAGGAAGCACCCATGCTCAAGAGCTAGG 1767
QY 1485 CGGCTCTTGGGCTTAGCACACCCAGATCCTTCTCACCCCTCCCTGGCTCCATGG 1544
Db 1768 GTGGCCCTTGAGTTACAGCACACCTAGTCCCTTCCCTCCCTGGCTC-----CC 1818
QY 1545 TGACCATGACTGAGGAGCAACTGGGCCACCGCTAGGTGGCCGCCAGAGCTGTTAATGACTT 1604
Db 1819 ATGACCATCTAGGAGCAACTGGGCCCAAGACAGATGCCAGAGCTGTTAATGGCT 1878
QY 1605 CAGCTGCTCACTCTCTCAAGATCAGCCTGTGGCATCTTTGCCCTTGGTGTCTGGCCACA 1664
Db 1879 CAGCTGCTCACTCTCTCAAGAGCAGCCTGTGGCATCTTTGCCCTTGGTGTCTGGCCACA 1935
QY 1665 GGGTCCAGGACTCTGGCCTTAGCCCAAGAGTGAAGAGAGCTTACA-GCGAGCTATGG 1723
Db 1936 CATGTGGTGTTCAGGGGACTCAGCCCTGAGGTGAAGGAGGATATCAGGAACAGCTATGG 1995
QY 1724 CAGCCCTGGGGCTTCCCT 1742
Db 1996 GAGCCCCAGGGGTTCCT 2014
```

```
RESULT 13
LOCUS MMU10102 1235 bp mRNA linear ROD 30-NOV-1995
DEFINITION Mus musculus bcl-x transmembrane deleted (bcl-x long) mRNA,
complete cds.
ACCESSION U10102
VERSION U10102.1 GI:506649
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1235)
AUTHORS Fang, W., Rivard, J.J., Mueller, D.L. and Behrens, T.W.
TITLE Cloning and molecular characterization of mouse bcl-x in B and T
lymphocytes
JOURNAL J. Immunol. 153 (10), 4388-4398 (1994)
MEDLINE 95052604
PUBMED 7963517
REFERENCE 2 (bases 1 to 1235)
AUTHORS Behrens, T.W.
TITLE Direct Submission
JOURNAL Submitted (26-MAY-1994) Timothy W. Behrens, Medicine, University of
Minnesota, 515 Delaware St. S.E., Minneapolis, MN 55455, USA
FEATURES
source
1..1235
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/cell_line="22D6"
/clone_lib="pre-B lymphocyte cDNA library"
/note="alternatively spliced transcript of bcl-x long,
GenBank Accession Number U10101"
1..1235
/gene="bcl-x long"
1..237
/gene="bcl-x long"
238..882
/gene="bcl-x long"
/standard_name="bcl-x tm-"
```

```
3'UTR
883..1235
/gene="bcl-x long"
ORIGIN
Query Match 51.3%; Score 893.4; DB 10; Length 1235;
Best Local Similarity 89.7%; Pred. No. 6e-248;
Matches 1022; Conservative 0; Mismatches 46; Indels 71; Gaps 3;
QY 1 CACAGAGCAGACCCAGTCAGTGAGCAGGTGTTTGGCAATGAGCTGTTGAGGCCATCT 60
Db 167 CACAGAGCAGACCCAGTAAGTGAGCAGGTGTTTGGCAATGAGCTGTTGAGGCCATCT 226
QY 61 CTATTATAAAATCTCTCAGAGCAACCGGGAGCTGGTGGTGTGCTTTCTCTCTACAAGC 120
Db 227 CTATTATAAAATCTCTCAGAGCAACCGGGAGCTGGTGGTGTGCTTTCTCTCTACAAGC 286
QY 121 TCTCCCAAGAGGATACAGCTGGAGTCAGTTTACGATGTCGAGAGCAAGGACTGAAG 180
Db 287 TTTCCCAAGAGGATACAGCTGGAGTCAGTTTACGATGTCGAGAGCAAGGACTGAAG 346
QY 181 CCCCAGAGAAACTGAAACCAAGAGGAGACCCCGAGTGCCATCAATGGCAACCCATCCT 240
Db 347 CCCCAGAGAAACTGAAACCAAGAGGAGGAGACCCCGAGTGCCATCAATGGCAACCCATCCT 406
QY 241 GGCACCTGGCGGATAGCCCGCGGTGAATGAGGCCACTGGCCACAGCAGCAGTTTGGATG 300
Db 407 GGCACCTGGCGGATAGCCCGCGGTGAATGAGGCCACTGGCCACAGCAGCAGTTTGGATG 466
QY 301 CGCGGAGGTAAATCCCATGSCAGCAGTGAAGCAAGCGCTGAGAGAGCTGGCGATGAGT 360
Db 467 CGCGGAGGTAAATCCCATGSCAGCAGTGAAGCAAGCGCTGAGAGAGCAGCGCATGAGT 526
QY 361 TTGAACCTGCGGTACCGGAGAGCATTTCAGTGATCAATCCAGCTTCATATAACCCAG 420
Db 527 TTGAACCTGCGGTACCGGAGAGCATTTCAGTGATCAATCCAGCTTCATATAACCCAG 586
QY 421 GGACAGCATATCAGAGCTTTGAAACAGGTAGTGAATGAATCTTTTCGGGATGGGTAAACT 480
Db 587 GGACCGGTATCAGAGCTTTGAGCAGGTAGTGAATGAATGAATCTTTTCGGGATGGGTAAACT 646
QY 481 GGGGTGCGATTGTGGCTCTTCTTCTTTCGGGGGCGACTGTGGTGAAGAGCGTAGACA 540
Db 647 GGGGTGCGATTGTGGCTCTTCTTCTTTCGGGGGCGACTGTGGTGAAGAGCGTAGACA 706
QY 541 AGGAGATGCAAGTATGCTGAGTCGATGCAAGTTGAGTGGCCACCTACCTGAATGACC 600
Db 707 AGGAGATGCAAGTATGCTGAGTCGATGCAAGTTGAGTGGCCACCTATCTGAATGACC 766
QY 601 ACCTAGAGCCTTGGATCCAGAGAAACCGCGCTCTGGGACACTTTTGTGTGATCTCTACGGGA 660
Db 767 ACCTAGAGCCTTGGATCCAGAGAAACCGCGCTCTGGGACACTTTTGTG----- 813
QY 661 ACAATGACAGCGCCAGAGACCGGAAAGGCCAGGAGCGTTTCAACCGCTGGTTCCTGACGG 720
Db 814 -----GACGG 818
QY 721 GCATGACTGTGGCTGGTGTAGTTCTCTGGGCTCACTCTTCACTGGAGTGAACAGACA 780
Db 819 GCATGACTGTGGCTGGTGTAGTTCTCTGGGCTCACTCTTCACTGGAGTGAACAGACA 878
QY 781 CTGACCGTCCACTCACTCTCACTCTCCACCTTCCCGCCACCAACAACCTCTCTTCAGCC 840
Db 879 CTGACCGTCCACTCACTCTCACTCTCCACCTTCCCGCCACCAACAACCTCTCTTCAGCC 938
QY 841 ACCATTGCTACAGGAGAACCACTACATGCAACTCACGCCCTTCCCTATTTATAGGTT 900
```

```
/codon_start=1
/product="bcl-x transmembrane deleted"
/protein_id="AA82174.1"
/db_xref="GI:506650"
/translation="MSQSNRELVDPLSYKLQKGYWSQFSDVEENRTEAPEETAE
RETPSALNGPNPWHLADSPAVNGATGHSISLDAREVIPAARKALREAGDEFLRYR
RAFSDLTSQHLITPGTAYQSFQVWVNFELFRDQVWGRIVAFPSFGGALCVBSFKEMQ
VLVSRASWWTATYLNHLEFWIQENGWDITFVDHDCGWCAGSLTLQSEVTRH"
883..1235
/gene="bcl-x long"
```

Db 939 ACCATTGCTACAGGAGAACCACTACATGCAACTCAGCCCTTGCCCTATCAGAGGTT 998  
Qy 901 GGGCTACAGGAGTCCCTGAGTACCTTCTAGATCTACACGCTTCTGTGAAGC 960  
Db 999 GGGCTACAGGAGTCCCTGCAATAGCTTCTAGGATCTACACACTTCTGTGAAGC 1058  
Qy 961 CACCTTCCCCCACATCTCAGTTCCCTTGGCTCAAAACTCACAAGGTTTTCCTCAGAT 1020  
Db 1059 CACCTTCCCCCACATCTCAGTTCTCTTGGCTCAAAACTCACAAGGTTTTCCTTAAC 1118  
Qy 1021 CAGCTCTTGGAGCTGCGAGAGTGGGAAGGGGTGTCTAGAGGGAAGAGCCTGCCT 1080  
Db 1119 CAGCTCTCTT-GAGGCTGCGAGAGCAGAGAGGGGTGTG--TGAGGGAGGAGCCTGTCT 1175  
Qy 1081 TGTTGGTGGGACCCGTATACCCCTGAGCCTCTCGGGAATGTTTTCGGCAGGAGCT 1139  
Db 1176 TGTTGGTGGGACCCCTAAATACCCCTGAGCCTCTCGGGAATGTTTTCAGCAGGAGCT 1234

RESULT 14  
LOCUS AF13327783 8022 bp DNA linear ROD 01-OCT-2002  
DEFINITION Mus musculus B-cell leukemia/lymphoma x (Bclx) gene, exon 3 and partial cds, alternatively spliced.  
ACCESSION AF133282  
VERSION AF133282.1 GI:13194516  
KEYWORDS 3 of 5  
SEGMENT Mus musculus (house mouse)  
SOURCE Mus musculus  
ORGANISM Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 8022)  
AUTHORS Yang,X.-F., Ye,Q., Press,B., Han,R.-Z., Bessing,C.H., Sleckman,B.P., Alt,F.W. and Cantor,H.  
TITLE Analysis of the complex genomic structure of Bcl-x and its relationship to Bcl-xgamma expression after CD28-dependent costimulation  
JOURNAL Mol. Immunol. 39 (1-2), 45-55 (2002)  
MEDLINE 22202614  
PUBMED 12213327  
REFERENCE 2 (bases 1 to 8022)  
AUTHORS Yang,X.-F. and Cantor,H.  
TITLE Direct Submission  
JOURNAL Submitted (08-MAR-1999) Cancer Immunology and AIDS, Dana-Farber Cancer Institute, Harvard Medical School, 44 Binney Street, Boston, MA 02115, USA

FEATURES  
source Location/Qualifiers  
1. 8022  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="129/SVJ"  
/db\_xref="taxon:10090"  
/chromosome="2"  
join(AFI33281.1:<1..517,571..>708)  
/gene="Bclx"  
/product="B-cell leukemia/lymphoma x"  
/note="alternative splicing exists in 3' untranslated regions resulting in multiple transcripts"  
join(AFI33281.1:<1..517,571..708)  
/gene="Bclx"  
/note="anti-apoptosis regulatory protein"  
/codon\_start=2  
/product="B-cell leukemia/lymphoma x"  
/protein\_id="AAK15455.1"  
/db\_xref="GI:13194520"  
/translation="LSQKGSQSDVEENRTEAPEETEARETSPSAINPSWHLA DSPAVNGATSSLSAREVIFPMIAVKQALRAGDEFELRYRRAFSDLTSQHLTPGT AYQSPQVNNLRFELGVNWRIVAFPSFGALCVESVDKMQVLVSRIASWMTYLNLD HLEPWIQNGWMDTFVDLYGNNAAESRKQGERFNRWFLTGMTVAGVVLLGSLFSRK" 571..>708  
/gene="Bclx"  
exon

ORIGIN  
Query Match 46.6%; Score 811.6; DB 10; Length 8022;  
Best Local Similarity 88.7%; Pred. No. 3.4e-224;  
Matches 992; Conservative 0; Mismatches 89; Indels 37; Gaps 9;  
/number=3  
Qy 633 TGGGACACTTTTGTGGGATCTCTACGGGAACAATGCAGCAGCGAGAGCGGAAAGGCCAG 692  
Db 568 TAGGACACTTTTGTGGGATCTCTACGGGAACAATGCAGCAGCGAGAGCGGAAAGGCCAG 627  
Qy 693 GAGCGTTTCAACCGCTGGTTCTGACGGGCATGACTGTGGCTGTGTAGTTTCTCTGGGC 752  
Db 628 GAGCGTTTCAACCGCTGGTTCTGACGGGCATGACTGTGGCTGTGTAGTTTCTCTGGGC 687  
Qy 753 TCACCTTTCAGTGGGAAGTACACAGACACTGACCGTCACTCAGCTCTCAGCTCCACCT 812  
Db 688 TCACCTTTCAGTGGGAAGTACACAGACACTGACCGTCACTCAGCTCTCAGCTCCACCC 747  
Qy 813 TGCCCCCACACCACTCTCTCTTCCAGCCACCATTTGTACAGGAGAACCACTACATGCAA 872  
Db 748 TGCCCCCACACCACTCTCTCTTCCAGCCACCATTTGTACAGGAGAACCACTACATGCAA 807  
Qy 873 CTACAGCCCTTCCCTTATATAGGGTTGGGCTTAGACGGAGTCCCTGCGAGTTAGCTTT 932  
Db 808 CTACAGCCCTTGGCTTATCAGAGGGTTGGGCTTAGACGGAGTCCCTGCGAATTAGCTTT 867  
Qy 933 CTAGAACTACACCGCTTCTGTGAAAGCACTTCCCTCCCACTCATGTTCCCTTGGCC 992  
Db 868 CTAGGATCTACCACTTCTGTGAAAGCACTTCCCTCCCACTCATGTTCTCTTGGCC 927  
Qy 993 TCRAAACTCAGAGGTTTTCCTCAGATCAGCTCTTGGAGGCTGGAGGAGTGGGAGG 1052  
Db 928 TCRAAACTCAGAGGTTTTCCTTAAACCAGCTCTTGGAGGCTGGAGGAGTGGGAGG 987  
Qy 1053 GGTGTGCTAGAGGAGGAGGAGCTGCTGTGTGTGGGAGCCCTGATTAACCTCTGAGCCTC 1112  
Db 988 GGTGTG--TGAGGGAGGAGGAGCTGCTGTGTGTGGGAGCCCTAATTAACCTCTGAGCCTC 1045  
Qy 1113 TCGGGAATGCTTTTCTGGCAGGAGCTGGAGAGCTCTTAAACCTCTTCCCTCCAGAGACT 1172  
Db 1046 TCGGGAATGCTTTTCCAGCAGGAGCT-----CTTCCCTCGAAGGAGACT 1090  
Qy 1173 AGATTGCTTGTGTTTTCATGCTGTGCTGCCTCAGAAATGATCCATTTCCTCTGCTGTG 1232  
Db 1091 AGCTTGTCTTGTGTTTTCATGCTGTGCTGCCTCAGAGTGAT--CATTTCCATCCCGCTGTG 1149  
Qy 1233 TCCCTGGGGGGCTTCTCTTCCCTTCCATCTCCACCCCTCCCTCCAGAGGAGTTCAGTGAG 1292  
Db 1150 TCTCTGGG-----TCTCCCTCTCTCCCTCCTCCACCCCTCCAGAGGAGTTCAGTGAG 1201  
Qy 1293 GTGCTTTTAGCCCTTTTGAATACTAAATATGAGGCTGTCTGGGATTAACAGAGGAGG- 1351  
Db 1202 GTGCTTTTAGCCCTTTTGAATACTAAATATGAGGCTGTCTGGGATTAATGAGACAAGGA 1261  
Qy 1352 -----ACCTCTCTCCCTGCTGTGGCTGGCC--AGACCCCTCACTCTGCTCTGAATGTTCT 1405  
Db 1262 CCAGAACTCTCTCTCCCTGCTGTGGCTGGCCAAACCTCCCTCTGCTCTGAATGTTCT 1321  
Qy 1406 CCTGAGGCTCTGCTAGAGTCCAGCCCTCCAGGAGGAGGAGCGAGCTGCGGAAAGT 1465  
Db 1322 CCTGAGGCTCTGCTAGAGTCTAGCCCTCCAGGAGGAGGCTTGTGCTGTAAGT 1381  
Qy 1466 CCACCTCTGAGAGCTGAGGGCTCTTGGGGCTTAGACCAACCCCTCAGATCCTTCTCCAC 1525  
Db 1382 CCACCTTGTAGACTAGGGTGGCTCTTTCAGCTTAGCACCCCTCCCAAGACTCTTGTTCAC 1441  
Qy 1526 CCCTCCCTGGCTCCATGGTGACCATGACTAGGAGGACCACTGGGCCCACTAGGTGCC 1585  
Db 1442 CCCTCCCTGGCTCC--CATGACCATGACTAGGAGACCA--TGGGCCCAAGATAGGTGCC 1498  
Qy 1586 CCAGAGCTGTTTAATGACTTTCAGTGCCTCACTTCTTCAAGATCAGCCTTGCGCATCTTT 1645

Db 1499 CCAGAGCTGTTAATGACTTACGCTCCTCACTTCTCGCAAGATTGGCTGTGGCATCTTT 1558  
 Qy 1646 GCCTTGGGTGCTGCCACAGGCTCCAGGACTCTGGCTTTAGCCCGAGAGTGAGAGGAAG 1705  
 Db 1559 GCCTTGGGTGCTGCCACAGGCTCCAGGACTCTGGCTTACGCCAGAGTGAGGGGAAG 1618  
 Qy 1706 CT-TACAGCGCAGCTATGAGGACCTTGGGGCTTCCCT 1742  
 Db 1619 CTATACAGAGCAGCTGTGGGAACCCAGGGGCTTCCCT 1656

RESULT 15  
 LOCUS AL731857 157220 bp DNA linear ROD 20-JUN-2002  
 DEFINITION Mouse DNA sequence from clone RP23-106A3 on chromosome 2, complete  
 sequence.

ACCESSION AL731857  
 VERSION AL731857.9 GI:21537530

KEYWORDS HTG.  
 SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 157220)  
 Almeida, J.

REFERENCE Direct Submission  
 AUTHORS Submitted (31-MAY-2002) Wellcome Trust Sanger Institute, Hinxton,  
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
 humquy@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk

TITLE On Jun 21, 2002 this sequence version replaced gi:21104246.  
 JOURNAL During sequence assembly data is compared from overlapping clones.  
 COMMENT Where differences are found these are annotated as variations  
 together with a note of the overlapping clone name. Note that the  
 variation annotation may not be found in the sequence submission  
 corresponding to the overlapping clone, as we submit sequences with  
 only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all  
 regions were either double-stranded or sequenced with an alternate  
 chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such  
 as compressions and repeats; all regions were covered by at least  
 one plasmid subclone or more than one M13 subclone; and the  
 assembly was confirmed by restriction digest. The following  
 abbreviations are used to associate primary accession numbers given  
 in the feature table with their source databases: Em, EMBL; Sw,  
 SWISSPROT; Tr, TrEMBL; Wp, WORMPEP; Information on the WORMPEP  
 database can be found at  
 http://www.sanger.ac.uk/Projects/C\_elegans/wormpep RP23-106A3 is  
 from the RPCI-23 Mouse PAC Library

constructed by the group of Pieter de Jong.  
 For further details see http://www.chori.org/bacpac/home.htm  
 VECTOR: pBAC3.6.

FEATURES  
 source Location/Qualifiers  
 1..157220  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 /chromosome="2"  
 /clone="RP23-106A3"  
 /clone\_lib="RPCI-23"

## ORIGIN

Query Match 46.6%; Score 811.6; DB 10; Length 157220;  
 Best Local Similarity 88.7%; Pred No. 3.2e-224;  
 Matches 992; Conservative 0; Mismatches 89; Indels 37; Gaps 9;  
 Qy 633 TGGGACACTTTTGTGATCTTACGGGAACAATGACGAGCGGAGCGGAAAGGCCAG 692  
 Db 82044 TAGGACACTTTTGTGATCTTACGGGAACAATGACGAGCGGAGCGGAAAGGCCAG 82103  
 Qy 693 GAGCGTTCAACCGCTGTTCTTACGCGGCATGACTGTGGCTGTGTAGTTCTGCTGGGC 752  
 Db 82104 GAGCGCTTCAACCGCTGTTCTTACGCGGCATGACTGTGGCTGTGTAGTTCTGCTGGGC 82163

Search completed: March 31, 2005, 02:53:54

Job time : 7406 secs

Qy 753 TCACCTTTCACTAGTCGGAAGTGAACAAGACACTGACCGTCCACTCACTCTCTCACTCTCCACCT 812  
 Db 82164 TCACCTTTCACTAGTCGGAAGTGAACAAGACACTGACCGTCCACTCACTCTCTCACTCTCCACCT 82223  
 Qy 813 TGCCCCCACCACCAACTCTCTCTTTCAGCACCACCATTTGTACAGGAGAACCACTACATGCAA 872  
 Db 82224 TGCCCCCACCACCAACTCTCTCTTTCAGCACCACCATTTGTACAGGAGAACCACTACATGCAA 82283  
 Qy 873 CTACAGCCCTTCCCTTATATAGGTTGGGCTTAGACGGAGTCCCTTCGAGTTAGCTTT 932  
 Db 82284 CTACAGCCCTTCCCTTATATAGGTTGGGCTTAGACGGAGTCCCTTCGAGTTAGCTTT 82343  
 Qy 933 CTAGAACTTACACGCTTCTGTGAAAGCACCCTTCCCTCCACATCTCAGTTCCCTTGCC 992  
 Db 82344 CTAGGATCTACCACTTCTGTGAAAGCACCCTTCCCTCCACATCTCAGTTCTCTTGCC 82403  
 Qy 993 TCRAAACTCAAGGTTTTTCTCAGATCAGCTCCTTGGAGGCTGGCAGGAGTGGGAAG 1052  
 Db 82404 TCRAAACTCAAGGTTTTTCTCAGATCAGCTCCTTGGAGGCTGGCAGGAGTGGGAAG 82463  
 Qy 1053 GGTGTGCTTAGAGGAGAGAGCCTTGTGTGGGACCTTGAATACCTCCCTGAGCCTC 1112  
 Db 82464 GGTGTG--TGAGGAGGAGAGCCTTGTGTGGGACCTTGAATACCTCCCTGAGCCTC 82521  
 Qy 1113 TCGGGAATGCTTTTCTGGCAGGAGGCTGGAGAGCTCTCTAAACCTCTTCCCTCCAGAGACT 1172  
 Db 82522 TCGGGAATGCTTTTCCAGCAGGAGGACT-----CTTCCCTTGAAGAGAGACT 82566  
 Qy 1173 AGATTGCTTGGTTTTGATGTGTGCGCTCAGAAATGATCCATTTCCATTTCTGCTGTG 1232  
 Db 82567 AGCTTGGCTTGGTTTTGATGTGTGCGCTCAGACGTGAT--CATTTCCATCCCGCTGTG 82625  
 Qy 1233 TCCCTGGGGGGCTCTCTTCCATCTCCACCCCTCCCTCCCTCCAGAGCCTTGAAGTGA 1292  
 Db 82626 TCTCTGGG-----TCTCCCTCTCTCCCTCTCCACCTCCAGAGCCTTGAAGTGA 82677  
 Qy 1293 GTGCTTTTAGCCCTTTTGAATAATAAATGAGGCTGTGTGGATTAACAGAGCAAGG- 1351  
 Db 82678 GTGCTTTTAGCCCTTTTGAATAATAAATGAGGCTGTGTGGATTAACAGAGCAAGG 82737  
 Qy 1352 -----ACCTCTCTCCCTCTGCTGGCTGGCC--AAGCCCTCCCTCTGCTGTAATGTTCT 1405  
 Db 82738 CCAGAACCTCTCTCCCTCTGCTGGCTGGCCAAACCTCTGCTGTAATGTTCT 82797  
 Qy 1406 CCTGAGGCTCTGCTGCTAGAGTCCAGCCCTCCAGAGGAGGAGCGAGCTGCGGAAAGT 1465  
 Db 82798 CCTGAGGCTCTGCTGCTAGAGTCCAGCCCTCCAGAGGAGGAGCGGCTTGAAGT 82857  
 Qy 1466 CCACCTTCGAGAGCCTGAGCGGCTCTTGGGGCTTAGCACACACCTCCAGATCTTCTCCAC 1525  
 Db 82858 CCACCTTCGAGAGCCTGAGCGGCTCTTGGAGCTTAGCACACCTCCAGATCTTCTCCAC 82917  
 Qy 1526 CCTCTCCCTGGCTCCATGGTGACCATGACTGAGGAGCAACTGGGCCCCACCTAGGTGCC 1585  
 Db 82918 CCTCTCCCTGGCTCC--CATGACCATGACTGAGGAGCAAA--TGGGCCCCAAGATAGGTGCC 82974  
 Qy 1586 CCAGAGCTGTTAATGACTTTCAGCTGCTCCTCTTGCAGAGTCCAGCTGAGGATCTTT 1645  
 Db 82975 CCAGAGCTGTTAATGACTTTCAGCTGCTCCTTCTGCAAGATTGGGCTGTGGCATCTTT 83034  
 Qy 1646 GCCTTGGGTGCTGGCCACAGGGTCCAGGACTCTGGGCTTAGCCCGAGAGTGAGAGGAAG 1705  
 Db 83035 GCCTTGGGTGCTGGCCACAGGGTCCAGGACTCTGGGCTTAGCCCGAGAGTGAGAGGAAG 83094  
 Qy 1706 CT-TACAGCGCAGCTATGGGAGCCCTGGGGCTTCCCT 1742  
 Db 83095 CTATACAGAGCAGCTGTGGGAACCCCGAGGGCTTCCCT 83132



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 31, 2005, 00:28:32 ; Search time 913 Seconds  
(without alignments)  
11294.841 Million cell updates/sec

Title: US-10-049-822A-1  
Perfect score: 1742  
Sequence: 1 cacagagcagaccagtgag.....ggagccctgggggttcctc 1742

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_16Dec04.\*

- 1: Geneseqn1980s.\*
- 2: Geneseqn1990s.\*
- 3: Geneseqn2000s.\*
- 4: Geneseqn2001as.\*
- 5: Geneseqn2001bs.\*
- 6: Geneseqn2002as.\*
- 7: Geneseqn2002bs.\*
- 8: Geneseqn2003as.\*
- 9: Geneseqn2003bs.\*
- 10: Geneseqn2003cs.\*
- 11: Geneseqn2003ds.\*
- 12: Geneseqn2004as.\*
- 13: Geneseqn2004bs.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1742	100.0	1742	4	Aaf75960 Rat wild-
2	1168	67.0	1466	10	Ad85177 Farnesyl
3	1038.2	59.6	2575	12	Ado19866 Human PRO
4	1038.2	59.6	2575	13	Acn40740 Tumour-as
5	1038.2	59.6	2575	13	Adp54991 Human PRO
6	1038.2	59.6	2575	13	Adp24509 PRO polyp
7	1038.2	59.6	2598	12	Adq97765 Human can
8	811.6	46.6	71594	12	Adq97761 Mouse can
9	787.8	45.2	863	12	Adh52632 Chinese h
10	741	42.5	926	2	Aag81698 Human thy
11	741	42.5	926	2	Aat40079 Bcl-xL ge
12	737.8	42.4	926	3	Aaz93614 Bcl-xL gen
13	737.8	42.4	926	4	Aas15189 Human bcl
14	737.8	42.4	926	4	Aac90810 Human bcl
15	737.8	42.4	926	6	Abk84766 Human cDN
16	737.8	42.4	926	8	Abt16641 Human bcl
17	737.8	42.4	926	10	Adt56779 Human bcl
18	737.8	42.4	926	10	Aad64187 Human bcl
19	737.8	42.4	926	11	Adi32104 Human cDN
20	737.8	42.4	926	12	Adh52630 Human ant

21	737.8	42.4	926	12	Ado19990
22	737.8	42.4	926	12	Adp13351
23	721.8	41.4	5903	13	Ado80674
24	721.8	41.4	6535	13	Ado80672
25	721.8	41.4	7257	13	Ado80673
26	703.6	40.4	7436	13	Ado80675
27	690.8	39.7	702	13	Ado80676
28	661.4	38.0	2386	10	Adg89403
29	661.4	38.0	2386	12	Adn04260
30	661.4	38.0	2386	13	Adp54628
31	661.4	38.0	2386	13	Adp23989
32	654	37.5	1472	12	Ado97762
33	637.4	36.6	7372	2	Aax33182
34	636	36.5	1748	10	Adb58615
35	636	36.5	1748	10	Adb53263
36	633.2	36.3	702	12	Adm45994
37	633.2	36.3	702	13	AdS74144
38	630	36.2	702	5	Aah43464
39	627.6	36.0	1236	5	Aas00247
40	599.2	34.4	1384	2	Aav17638
41	581	33.4	739	12	Adg65218
42	581	33.4	747	4	Aaf30926
43	579.4	33.3	747	12	Adg65209
44	576.8	33.1	79544	12	Ado97764
45	572	32.8	695	6	Abt09346

## ALIGNMENTS

### RESULT 1

AAf75960	AAf75960 standard; cDNA; 1742 BP.
ID	AAf75960 standard; cDNA; 1742 BP.
XX	XX
AC	AAf75960;
XX	XX
DT	22-MAY-2001 (first entry)
XX	XX
DE	Rat wild-type Bcl-xL cDNA.
XX	XX
KW	Rat Bcl-xL; apoptosis inhibitor; programmed cell death inhibitor;
KW	wild-type; antiapoptotic; cell death-associated disease;
KW	tissue transplant preservative; ss.
XX	XX
OS	Rattus norvegicus.
XX	XX
PN	WO200112807-A1.
XX	XX
PD	22-FEB-2001.
XX	XX
PF	17-AUG-2000; 2000WO-JP005502.
XX	XX
PR	17-AUG-1999; 99JP-00230642.
XX	XX
FA	(NISC-) JAPAN SCI & TECHNOLOGY CORP.
PI	Ohta S, Asoh S;
XX	XX
DR	WPI; 2001-211219/21.
XX	XX
PT	P-FSDB; AAB73303.
XX	XX
PT	Modified cDNA of rat bcl-x gene and encoded protein with membrane permeability to enhance uptake for effective inhibition of cell death e.g. apoptosis, useful in remedies for diseases associated with cell death.
XX	XX
PS	Claim 1; Page 43-45; 56pp; Japanese.
XX	XX
CC	The invention relates to a mutant rat Bcl-x protein and the cDNA encoding it. The mutant rat Bcl-x protein (Bcl-xFNK) has the substitutions Y22F, Q26N, and R155K relative to the wild-type Bcl-xL protein. The invention also encompasses recombinant vectors and host cells comprising the modified nucleic acid sequence. The mutant Bcl-x protein is able to



XX OS	Homo sapiens.	Db	452	GGACCGCGTATCAGAGCTTTGAGCAGGCTAGTGAATGAATCTCTTCGGGATGGAGTAAACT	511
XX PN	W02003038129-A2.	Qy	481	GGGCTCGCATTTGTGGCTCTTCTTCCTTTGGCGGGGCACTGTGCTGTGAAAGCCTAGACA	540
XX XX		Db	512	GGGCTCGCATCGTGGCTTTTCTCTCTTTGGCGGGGCACTGTGCTGTGAAAGCCTAGACA	571
XX PD	08-MAY-2003.	Qy	541	AGGAGATGACAGGATTTGGTGTGATTCGGAATTCGAAGTTGGATGGCCACCTACCTGAATGACC	600
XX PF	30-OCT-2002; 2002WO-US034784.	Db	572	AGGAGATGACAGGATTTGGTGTGATTCGGAATTCGAAGTTGGATGGCCACCTATCTGAATGACC	631
XX PR	30-OCT-2001; 2001US-0338997P.	Qy	601	ACCTAGAGCCTTTGGATCCAGGAGAACGGCGCTGGGACACTTTTGTGGATCTCTACCGGA	660
XX PR	30-OCT-2001; 2001US-0340938P.	Db	632	ACCTAGAGCCTTTGGATCCAGGAGAACGGCGCTGGGACACTTTTGTGGATCTCTACCGGA	691
XX PR	30-OCT-2001; 2001US-0341012P.	Qy	661	ACAAATGACAGCAGCCGAGAGCCGGAAGCCAGGAGCGTTTCAACCGCTGGTTCCTGACGG	720
XX PA	(ORTH ) ORTHO CLINICAL DIAGNOSTICS INC.	Db	692	ACAAATGACAGCAGCCGAGAGCCGGAAGCCAGGAGCGTTTCAACCGCTGGTTCCTGACGG	751
XX PI	Raponi M;	Qy	721	GCATGACTGTGGCTGGTGTAGTTCTGTCTGGGCTCACTCTTCAGTCGGAGTGCACGACA	780
XX DR	WPI; 2003-513497/48.	Db	752	GCATGACTGTGGCTGGTGTAGTTCTGTCTGGGCTCACTCTTCAGTCGGAGTGCACGACA	811
XX PT	Determining whether a patient will respond to treatment with a farnesyl	Qy	781	CTGACCGTCCACTCACCTCTCACCTCCACCTTCCACCTTCCACCTTCCACCTTCCACCT	840
XX PT	transferase inhibitor, by analyzing the expression of gene that is	Db	812	CTGACCGTCCACTCACCTCTCACCTCCACCTTCCACCTTCCACCTTCCACCTTCCACCT	871
XX PT	differentially modulated in the presence of the inhibitor.	Qy	841	ACCAATGCTACAGGAGAACCACTACATGCAACTCAGCCCTTCCCTTATATAGGGTT	900
XX PS	Disclosure; SEQ ID NO 396; 346pp; English.	Db	872	ACCAATGCTACAGGAGAACCACTACATGCAACTCAGCCCTTCCCTTATATAGGGTT	931
XX CC	The invention relates to a method of determining whether a patient will	Qy	901	GGGCTAGACGAGTCCCTGCACTTGTAGTCTTCTAGTAACTACACGCTTCTGTGAAAGC	960
XX CC	respond to treatment with a farnesyl transferase inhibitor (FTI), by	Db	932	GGGCTAGACGAGTCCCTGCACTTGTAGTCTTCTAGTAACTACACGCTTCTGTGAAAGC	991
XX CC	analyzing the expression of gene that is differentially modulated in the	Qy	961	CACCTTCCCTCCACCTCTCAGTCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCT	1020
XX CC	presence of an FTI. The method is useful for determining whether a	Db	992	CACCTTCCCTCCACCTCTCAGTCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCT	1051
XX CC	patient will respond to treatment with a FTI such as (B)-6-(farnilo(4-	Qy	1021	CAGCTCTCTTGAGGCTGGCAGGAGTGGGAGGAGTGTCTAGAGGGAGGAGAGCTGCT	1080
XX CC	chlorophenyl)(1-methyl-1H-imidazol-5-yl)methyl)-4-(3-chlorophenyl)-1-	Db	1052	CAGCTCTCTTGAGGCTGGCAGGAGTGGGAGGAGTGTCTAGAGGGAGGAGAGCTGCT	1108
XX CC	methyl-2-(1H)quinolinone, monitoring the therapy of a patient, treating a	Qy	1081	TGTTGTGGGAGCCTGATTAACCTGAGCC-TCTCGGGAATGCTTTTCTGGAGGAGCT	1139
XX CC	patient with leukemia with FTI if the analysis indicates that the patient	Db	1109	TGTTGTGGGAGCCTGATTAACCTGAGCC-TCTCGGGAATGCTTTTCTGGAGGAGCT	1168
XX CC	will respond. This sequence corresponds to a gene whose expression may be	Qy	1140	GGAGAGCTCTCTAAACCTCTTCCCTCCAGAGACTAGATTCCTTGGTGTGTGTGG	1199
XX CC	modulated in the presence of FTI.	Db	1169	-----CTTCCCTTGAAGAGAGTGTCTTGGTGTGTGTGGTGTGTGG	1213
XX SQ	Sequence 1466 BP; 332 A; 402 C; 404 G; 328 T; 0 U; 0 Other;	Qy	1200	CCTCAGAAATGATTCATTTCCCATTTCTGTCTGTCTTCCCTGGGGCGGCTCCTCTTCCCATC	1259
	Query Match 67.0%; Score 1168; DB 10; Length 1466;	Db	1214	CCTCAGAGGTATC-ATTTCCCATCGCTGTCTTCCCTGGG-----TCTCCCTCTC	1263
	Best Local Similarity 91.0%; Pred. No. 0;	Qy	1260	TCCACCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT	1319
	Matches 1331; Conservative 0; Mismatches 95; Indels 36; Gaps 7;	Db	1264	CCCATCTCCACCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT	1323
Qy	1 CACAGAGCAGACCCAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT	Qy	1320	AAATGACGCTGTCTGGGATTAACGAGGCAAG-----ACCTCTCTCCCTCCCTCTGGGCT	1373
Db	32 CACAGAGCAGACCCAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT	Db	1324	AAATTCAGGCTGTCTGGGATTAACGAGGCAAG-----ACCTCTCTCCCTCCCTCTGGGCT	1383
Qy	61 CTATTATATAAATGTCTCAGAGCAACCGGAGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT	Qy	1374	GGCC- AAGCCCCCACTCTCTGGTCTGAATGTTCTCTGAGGCTCTCTGGCTAGAGTCCAGCC	1432
Db	92 CTATTATATAAATGTCTCAGAGCAACCGGAGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT	Db	1384	GGCCAAACCCCACTCTCTGGTCTGAATGTTCTCTGAGGCTCTCTGGCTAGAGTCCAGCC	1443
Qy	121 TCTCCAGAAAGGATACAGCTGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT	Qy	1433	CCACCCAGGAGGAGGAGGAG 1454	
Db	152 TTTCCAGAAAGGATACAGCTGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT	Db	1444	TCACCCAGGAGGAGGAGGAG 1465	
Qy	181 CCCAGAGAAACTGAACCCAGAAAGGGAGACCCCGAGTGCCATCAATGGCAACCCATCCT				
Db	212 CCCAGAGAAACTGAACCCAGAAAGGGAGACCCCGAGTGCCATCAATGGCAACCCATCCT				
Qy	241 GGCACTCGCGGATAGCCCGGCTGAATGAGGACACTGGCCACAGCAGCAGTGTGGATG				
Db	272 GGCACTCGCGGATAGCCCGGCTGAATGAGGACACTGGCCACAGCAGCAGTGTGGATG				
Qy	301 CGCGGGAGTATCCCATGGCAGCAGTGAAGCAAGCCGCTGAGAGAGCTGGCGATGAGT				
Db	332 CGCGGGAGTATCCCATGGCAGCAGTGAAGCAAGCCGCTGAGAGAGCTGGCGATGAGT				
Qy	361 TTGAACCTCGGTACCGGAGAGCTTCAGTGTATTAACATCCAGCTTCCATATAACCCAG				
Db	392 TTGAACCTCGGTACCGGAGAGCTTCAGTGTATTAACATCCAGCTTCCATATAACCCAG				
Qy	421 GGACAGCATATCAGAGCTTTGAAACAGGCTAGTGAATGAATCTTTTCGGGATGGGTAATCT				







601 ACCTAGAGCCTTGGATCCAGAGAAACGGCGCTGGGACACTTTTGTGTGATCTCTACGGGA 660  
 896 ACCTAGAGCCTTGGATCCAGAGAAACGGCGCTGGGATACCTTTTGTGGAATCTCTATGGGA 955  
 661 ACAATGACAGCAGCCGAGAGCCGGAAGACCGAGAGCGTTTCAACCGCTGGTTCCTGACGG 720  
 956 ACAATGACAGCAGCCGAGAGCCGGAAGACCGAGAGCGTTTCAACCGCTGGTTCCTGACGG 1015  
 721 GCATGACTGTGGCTGGTGTAGTTCTGTCTGGGCTCACTCTTCAGTCGGAAGTGCACAGACA 780  
 1016 GCATGACTGTGGCTGGTGTAGTTCTGTCTGGGCTCACTCTTCAGTCGGAATGCACAGACA 1075  
 781 CTGACCGTCCACTCACTCTCACTCCACCTTGCCTCCACACACAACTCTCTCTTCAGCC 840  
 1076 CTGACCATCCACTCACTCTCCACCTCCACCTCTCTCTCTCCACATCTCTCCGTCAGCC 1135  
 841 ACCATTGCTACAGAGAACCACTACATGCAACTCAGCCCTCTCTCCCTATTTATAGGGTT 900  
 1136 GCCATTGCCACAGAGAACCACTACATGCAAGCCATGCCCTGCCATCAGAGGTT 1195  
 901 GGGCTAGACGAGTCCCTCGAGTCTAGTCTTCTAGATCTACCAAGCTTCTGTGAAGC 960  
 1196 GGGCCAGATCTGGTCCCTCGAGTCTAGTCTTCTAGATTTATCACACTCTCTGTGAGA-- 1253  
 961 CACCTTCCCCCAATCTCAGTCTCCCTGGCTCAAACTCACAAGGTTTCTCTCAGAT 1020  
 1254 -----CCCCCACTCAGTCTCCCTGGCTCAGAACTCAGAA--AAATTCACAAAT 1305  
 1021 CAGCTCTCTGAGGCTGCGAGAGTGGGAAGGGGTGCTAGAGGG--AGAAGAGCCTGC 1078  
 1306 CTGTCCAAAGAGGCTGCGAGATGGAAGGGTGTGCTGGGCGAGGAGGCGCCTAC 1365  
 1079 CTGTGTGTGGAGCCCTGATTAACCTCAGCTCTCGGGAATGCTTCTGGCAGGAGC 1138  
 1366 CTGATTGTGTCAACCT--TACCCCTAGCCTCCCTGAAATGTTTTCTGCCAGGAGC 1423  
 1139 TGGAGAGCTCTAAACCTCTTCCC----CAGAGACTAGATTGCTTGGTGTGATGTG 1194  
 1424 TTGAAGTTTTCAGAACTCTTCCCAGAAAGGAGCTAGATTGCTTGTGTTGATGTT 1483  
 1195 TGTGGCTCTCAGAAATGATCCATTTCCCAATCTGCTGTGCTCCCTGGGCGGCTCCTCCTTC 1254  
 1484 TGTGGCTCTCAGAAATGATCAATTTTCCCTCCCACTCTCCCACTAACTGCGTTCCTCTT 1543  
 1255 CCATCTCCACCCCTCCCTCCAGAGCCATTTGAGTGGTGTGTTAGCCCTTTTGACTA 1314  
 1544 CC--TTCCATCCCTACCCCTTAAGAGCCATTTAGGGGCACTTT-----TG 1587  
 1315 ACTAAATGAGGCTGTGGGATAACGAGGCAAGGACCT-----CCTCCCACTGT 1368  
 1588 ACTAGGATTCAGGCTGTGGGATAAAGATGCAAGGACAGGACTCCTCTCCTCCTCT 1647  
 1369 GGCCTGGGCA-AGCCCTCCTCTGCTGATGTTCTCTGAGGCTCTGCTAGAGTC 1427  
 1648 GGCCTGCTAGATCTCTCACTCCCACTGCTCCCAAGAGCTCTGCTAGAGGC 1707  
 1428 CAGCCCAACCAAGAG--GAGGAGCGAGCTGCGGAAGTCCACCTTCGAGAGCCTGA 1484  
 1708 CAGCCCAACCAAGAGGAGGAGGCTGCTATAGCTACAGGAAGCACCCTATGCTCAAGCTAGG 1767  
 1485 GCGGCTCTGGGCTTAGCACACCCAGATCTCTTCTCCACCTCCCTGCTCCATGG 1544  
 1768 GTGGGCTTGCAGTTTCAGACACCTTAGTCTCTTCCCTCTCTGCTGCTC-----CC 1818  
 1545 TGACCATGACTGAGGACCAATGCGGCCCAAGCTAGGTGCCCCAGAGCTGTTAATGACTT 1604  
 1819 ATGACCATGAGGACCAATGCGGCCCAAGAGAGATGCCCCAGAGCTGTTATGGCT 1878  
 1605 CAGTGCCTCACTCTCTGAGATCAGCTGTGGCATCTTTGCTTGGTGTGCTGCGCA 1664  
 1879 CAGTGCCTCACTCTCTGAGAGAGAGCTGTGGCATCTTTGCTTGGG-----CTGCTCT 1935  
 1665 GGGTCCAGGAGCTCTGGCTTAGCCCAAGTGAAGGAGGAGCTTACA-GCGCAGCTATGG 1723

Db 1936 CATGTTGGTTTCAGGGGACTCAGCCCTGAGTGAAGGAGCTATCAGGAACAGCTATGG 1995  
 Qy 1724 GAGCCCTGGGGCTTCCCT 1742  
 Db 1996 GAGCCCAAGGCTTCCCT 2014  
 RESULT 5  
 ADP54991  
 ID ADP54991 standard; cDNA; 2575 BP.  
 XX AC ADP54991;  
 XX 18-NOV-2004 (first entry)  
 XX Human PRO cDNA sequence SEQ ID NO:967.  
 DE human; PRO; immune related disease; inflammatory immune response;  
 XX immune response stimulation; antiallergic; antianaemic; antiarthritic;  
 XX antiasthmatic; antidiabetic; antinflammatory; antipsoriatic;  
 XX antineuritic; antithyroid; CNS; dermatological; gastrointestinal;  
 XX haemostatic; hepatotropic; immunostimulant; immunosuppressive; muscular;  
 XX nephrotropic; neuroprotective; osteopathic; respiratory; vasotropic;  
 XX virucide; gene therapy; gene; ss.  
 OS Homo sapiens.  
 XX WO2004039956-A2.  
 XX 13-MAY-2004.  
 XX 28-OCT-2003; 2003WO-US034381.  
 XX 29-OCT-2002; 2002US-0422472P.  
 XX (GETH ) GENENTECH INC.  
 XX Aggarwal S, Clark H, Gurney AL, Schoenfeld J, Williams PM;  
 FI Wood WI, Wu TD;  
 XX WPI; 2004-376182/35.  
 XX P-PSDB; ADP54992.  
 XX New PRO polynucleotides and polypeptides, useful in useful in diagnosing  
 PT and treating an immune related disease, e.g. systemic lupus  
 PT erythematosus, rheumatoid arthritis, diabetes mellitus or asthma and in  
 PT stimulating an immune response.  
 XX Claim 2; SEQ ID NO 967; 3009pp; English.  
 XX The present invention describes an isolated PRO nucleic acid (1). Also  
 CC described: (1) a vector comprising (1); (2) a host cell comprising the  
 CC vector of (1); (3) a process for producing a PRO polypeptide; (4) an  
 CC isolated PRO polypeptide; (5) a chimeric molecule comprising the  
 CC polypeptide of (4) fused to a heterologous amino acid sequence; (6) an  
 CC antibody which specifically binds to a polypeptide of (4); (7) a  
 CC composition of matter comprising a polypeptide of (4), an agonist or  
 CC antagonist of the polypeptide or an antibody that binds to the  
 CC polypeptide in combination with a carrier; (8) an article of manufacture  
 CC comprising a container, a label on the container and a composition of  
 CC matter of (7); (9) a method of treating an immune related disease in a  
 CC mammal; (10) a method for determining the presence of a PRO polypeptide  
 CC in a sample suspected of having the polypeptide; (11) a method of  
 CC diagnosing an immune related disease or an inflammatory immune response  
 CC in mammal; (12) a method of identifying a compound that inhibits or  
 CC mimics the activity of or expression of a gene encoding a PRO polypeptide  
 CC ; and (13) a method of stimulating the immune response in a mammal. The  
 CC PRO sequences have antiallergic, antianaemic, antiarthritic,  
 CC antineuritic, antidiabetic, antinflammatory, antipsoriatic,  
 CC antirheumatic, antithyroid, CNS, dermatological, gastrointestinal,  
 CC haemostatic, hepatotropic, immunostimulant, immunosuppressive, muscular,  
 CC nephrotropic, neuroprotective, osteopathic, respiratory, vasotropic and









	Db	60568	TCTCTGGG-----TCTCCCTCCTCCCACATCTCACCCCCCAAGAGCCATTGTAGTGAG	60619
	Qy	1293	GTGCTTTTAGCCCTTTTGACTAACTAAAAAATGCAGGCTGCTTTGGGATAACGAGGCAAGG-	1351
	Db	60620	GTGCTTTTAGCCCTTTTGACTAATTAGAAATTCAGGCTGCTTTGGGATAATGAGACAAGGA	60679
	Qy	1352	----ACCCTCTCCCAACCTGTGGCTCTGGCC--AAGCCCCCACTCTCGTGTGTAATGTCTCT	1405
	Db	60680	CCAAGAACCCTCTCCCCACCTGTGGCTGTGCCAAAACCCCCCATCTGTGTGTCGAATGTCTCT	60739
	Qy	1406	CCTGAGGCCCTCTGGCTAGAGTCCAGCCCAACCCAGAGGAGGAGCGAGCTCGCGAAAAGT	1465
	Db	60740	CCTGAGGCCCTCTGGCTTAGAGCTTAGCCCCAACCCAGGGGAGGGGCTTAGCTGCTGAAGT	60799
	Qy	1466	CCACCTTCGAGAGCCTGAGCGGCTCTTGCGGCTTAGCACACCACCACAGATCCTCTCCAC	1525
	Db	60800	CCACCTTGCTAGAGCTAGGGTGCTCTTGACGCTTAGCACCCCCCAAGATCCTCTGTCCAC	60859
	Qy	1526	CCCTCCCTGGCTCCATGTGTGACCATGTAGTGGGACCAAATGGGGCCACCGCTAGGTGCC	1585
	Db	60860	CCCTCCCTGGCTCC--CATGACCATGTAGTGGGACCAA--TGGGCCCAAGATAGGTGCC	60916
	Qy	1586	CCAGAGCTGTTAATGACTTACGTGCTCCTCACTTCCTGCAAGATCAGCCTGTGGCATCTTT	1645
	Db	60917	CCAGAGCTGTTAATGACTTACGTGCTCCTCACTTCCTGCAAGANTGGCCCTGTGGCATCTTT	60976
	Qy	1646	GCCTTGGGTGCTGGCCACAGGGTCCAGGAGACTCTGGCCTTAGGCCACAGAGTGAGAGGAAG	1705
	Db	60977	GCCTTGGGTGTTGGCCACAGGGTCAAGGAGACTCTGGCCTCAGCCCCAGAGTGAGGGGAAG	61036
	Qy	1706	CT--TACAGCCGACTTAGGAGCCCTGGGGGCTTCCTCT	1742
	Db	61037	CTATACAGACAGCTGTGGGAAACCCAGGGGCTTCCTCT	61074

## RESULT 9

ADH52632	
ID	ADH52632 standard; DNA; 863 BP.
XX	
XX	
AC	ADH52632;
XX	
DT	25-MAR-2004 (first entry)
XX	
XX	
DE	Chinese hamster anti-apoptosis bcl-xL wild-type DNA.
XX	
KW	mammalian myeloma host cell; protein production; anti-apoptosis;
KW	cell death; Chinese hamster; bcl-xL; wild-type; ds; gene.
XX	
OS	Cricetulus griseus.
XX	
PN	US2003219871-A1.
PD	
XX	27-NOV-2003.
XX	
XX	
PF	28-MAR-2003; 2003US-00402017.
XX	
XX	
PR	02-APR-2002; 2002US-0369307P.
XX	
PA	(BOEH ) BOEHRINGER INGELHEIM PHARMA GMBH & CO KG.
XX	
PI	Enenkel B, Meents H, Fusseneegger M;
XX	
XX	
DR	WPI; 2004-033642/03.
DR	P-PSDB; ADH52633.
XX	
XX	
PT	New genetically engineered hamster or murine myeloma host cells
PT	comprising enhanced levels of active anti-apoptosis genes, useful for
PT	producing complex protein therapeutics.
XX	
PS	Claim 57; SEQ ID NO 3; 46pp; English.
XX	
CC	The invention relates to a novel mammalian host cell for producing











Db 184 TTTCCAGAGGATACAGCTGGAGTCACTTTAGTGTGTAAGAGAACAGGACTGAGG 243  
Qy 181 CCCAGAGAAAGTGAACAGAGAGGAGAGACCCAGTGGCCATCAATGGCAACCCATCCT 240  
Db 244 CCCAGAGAGGACTGAATCGAGATGGAGACCCAGTGGCCATCAATGGCAACCCATCCT 303  
Qy 241 GGCACCTGGCGATAGACCCCGGCGTGAATGAGACCACTGGCCACAGCAGAGTTGGATG 300  
Db 304 GGCACCTGGCAGACAGACCCCGGCGTGAATGAGACCACTGGCCACAGCAGAGTTGGATG 363  
Qy 301 CGCGGAGGTAATCCCATGACAGCAGTGAAGAGCGCTGAGAGAGCTGCGATAGT 360  
Db 364 CGCGGAGGTAATCCCATGACAGCAGTGAAGAGCGCTGAGAGAGCTGCGATAGT 423  
Qy 361 TTGAACCTGGCTACCGGAGACATTCAGTGTATTAACATCCAGCTTCATATAACCCAG 420  
Db 424 TTGAACCTGGCTACCGGAGACATTCAGTGTATTAACATCCAGCTTCATATAACCCAG 483  
Qy 421 GGACAGCATATCAGAGCTTTGAACAGGTAGTGAATGAATCTTTTCGGGATGGGTAACT 480  
Db 484 GGACAGCATATCAGAGCTTTGAACAGGTAGTGAATGAATCTTTTCGGGATGGGTAACT 543  
Qy 481 GGGTGGCATTTGGCCCTTTCTCTCTTTGGGGGCACTGCGTGAAGAGCGTAGACA 540  
Db 544 GGGTGGCATTTGGCCCTTTCTCTCTTTGGGGGCACTGCGTGAAGAGCGTAGACA 603  
Qy 541 AGGAGATCAGAGTATTTGTGAGTGGATTCGAAGTTGGATGGCCACCTACCTGAATGACC 600  
Db 604 AGGAGATCAGAGTATTTGTGAGTGGATTCGAAGTTGGATGGCCACCTACCTGAATGACC 663  
Qy 601 ACCTAGAGCTTGGATCCAGAGAACCGCGCTGGGATACCTTTGTGGAACTCTATGGGA 723  
Db 661 ACATGACAGCGGAGAGCGGAGAACCGCAGGCGTTTCAACCGCTGTTCTGACGG 720  
Db 724 ACATGACAGCGGAGAGCGGAGAACCGCAGGCGTTTCAACCGCTGTTCTGACGG 783  
Qy 721 GCATGACTGTGGTGTGTAGTTCTGTGGGTCACTCTTCACTCGGAAGTGACAGACA 780  
Db 784 GCATGACTGTGGTGTGTAGTTCTGTGGGTCACTCTTCACTCGGAAGTGACAGACA 843  
Qy 781 CTGACCGTCCACTCACTTCACTTCCACCTTGGCCACCCACCACTCTCTTTCAGCC 840  
Db 844 CTGACCGTCCACTCACTTCACTTCCACCTTGGCCACCCACCACTCTCTTTCAGCC 903  
Qy 841 ACCATTGCTACAGGAGAAC 861  
Db 904 GCCATTGCCACGAGAGAAC 924

## RESULT 14

AAC90810  
ID AAC90810 standard; cDNA; 926 BP.

XX AAC90810;

XX AAC90810;

DT 16-MAR-2001 (first entry)

DE Human Bcl-xL nucleotide sequence SEQ ID NO:3.

XX Human; Bcl-2; Bcl-xL; Bax; VDAC; apoptosis inhibitor; detection;  
KW apoptosis promoter; diagnosis; ss.

XX Homo sapiens.

XX JP2000287689-A.

XX 17-OCT-2000.

XX 08-APR-1999; 99JP-00101888.

XX

PR 08-APR-1999; 99JP-00101888.

XX (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.

XX WPI; 2001-065575/08.

XX P-PSDB; AAB50538.

XX Screening of an apoptosis inhibitor or promoter which can be used as a  
PT drug and a diagnostic agent for various diseases caused by apoptosis  
PT inhibition or apoptosis promotion.

XX Claim 16; Page 14-15; 22pp; Japanese.

XX The present invention describes a method for screening for an apoptosis  
CC inhibitor or an apoptosis promoter in which VDAC-liposome, an index  
CC substance which can pass VDAC and a sample are incubated and the change  
CC in the concentration of the index substance during the incubation is  
CC detected to judge the presence of apoptosis inhibition or apoptosis  
CC promotion. The apoptosis inhibitor or the apoptosis promoter can be used  
CC as a drug and a diagnostic agent for various diseases caused by apoptosis  
CC inhibition or apoptosis promotion. The present sequence encodes the human  
CC Bcl-xL protein, which is an apoptosis inhibitor used in the  
CC exemplification of the present invention

XX Sequence 926 BP; 220 A; 249 C; 264 G; 193 T; 0 U; 0 Other;

Query Match 42.4%; Score 737.8; DB 4; Length 926;

Best Local Similarity 91.1%; Pred. No. 5.4e-202;

Matches 784; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

Qy 1 CACAGAGCAGACCAGTGTAGTGAGCAGGTGTTTGGACAATGGACTGTTGAGCCCATCT 60

Db 64 CAGAGACGAGACTCAGTGTAGTGAGCAGGTGTTTGGACAATGGACTGTTGAGCCCATCC 123

Qy 61 CTATTATATAAATGTCTCAGAGCAACCGGAGCTGGTGGTGTGACTTCTCTCTTACAAGC 120

Db 124 CTATTATATAAATGTCTCAGAGCAACCGGAGCTGGTGGTGTGACTTCTCTCTTACAAGC 183

Qy 121 TCTCCCAAGAAAGGATACAGCTGGAGTCACTTTAGCGATGTCTGAAGAGAACAGGACTGAAG 180

Db 184 TTTCCCAAGAAAGGATACAGCTGGAGTCACTTTAGCGATGTCTGAAGAGAACAGGACTGAAG 243

Qy 181 CCCCAGAGAAAGTGAACAGAGAGGAGACCCCGAGTGGCCATCAATGGCAACCCATCCT 240

Db 244 CCCCAGAGAGGACTGAATCGAGATGGAGACCCCGAGTGGCCATCAATGGCAACCCATCCT 303

Qy 241 GGCACCTGGCGATAGCCCGCGTGAATGAGCCACTGGCCACAGCAGCAGTTTGGATG 300

Db 304 GGCACCTGGCGATAGCCCGCGTGAATGAGCCACTGGCCACAGCAGCAGTTTGGATG 363

Qy 301 CGCGGAGGTAATCCCATGGCAGCTGAAGAGCGCTGAGAGAGCTGCGCATGAGT 360

Db 364 CC CGGAGGTGATCCCATGGCAGCTGAAGAGCGCTGAGAGAGCTGCGCATGAGT 423

Qy 361 TTGAACCTGGGTACCGGAGAGCATTCAGTGAATCTAAACATCCAGCTTCATATAACCCAG 420

Db 424 TTGAACCTGGGTACCGGCGGCATTCAGTGAATCTAAACATCCAGCTTCATATAACCCAG 483

Qy 421 GGACAGCATATCAGAGCTTTGAACAGGTAGTGAATGAATCTTTTCGGGATGGGTAACT 480

Db 484 GGACAGCATATCAGAGCTTTGAACAGGTAGTGAATGAATCTTTTCGGGATGGGTAACT 543

Qy 481 GGGGTGCGCATTTGGCCCTTTCTTCTTTTGGCGGGGCACTGTGGGTGAAAGCGTAGACA 540

Db 544 GGGGTGCGCATTTGGCCCTTTCTTCTTTTGGCGGGGCACTGTGGGTGAAAGCGTAGACA 603

Qy 541 AGGAGATGCAAGTATTTGGTGAATCGGATTCGAATGGATGGCCACCTTACCTGAATGACC 600

Db 604 AGGAGATGCAAGTATTTGGTGAATCGGATTCGAATGGATGGCCACCTTACCTGAATGACC 663

Qy 601 ACCTAGAGCTTGGATCCAGGAGAACCGCGCTGGGACACTTTTGTGTGATCTCTACGGGA 660

Db 664 ACCTAGAGCTTGGATCCAGGAGAACCGCGCTGGGATACCTTTTGTGTGATCTCTACGGGA 723



Db	724		ACATG	CAGC	CGAG	CGCG	CGAAG	GGCC	AGGA	CGCTT	CAAC	CGCT	GGTCT	CTG	ACGG	783	
Qy	721	GCATG	ACTG	TGGC	TGGT	GTAGT	CTGCT	GGGCT	CAC	TCTT	CAGT	CGGA	AGT	GAC	CAGACA	780	
Db	784	GCATG	ACTG	TGGC	CGGG	TGGTCT	TGCT	GGGCT	CAC	TCTT	CAGT	CGGA	AAAT	GAC	CAGACA	843	
Qy	781	CTGAC	CGT	TCC	ACT	CAC	TCT	CAC	CTT	CGCC	CCCA	CCAC	CA	ACTCT	CTCTT	CGCC	840
Db	844	CTGAC	CAAT	CCACT	CTAC	CCCT	CCCC	ACCC	CCCTT	CTCT	GTGCT	C	CA	CCAT	CCTCGT	CCAGCC	903
Qy	841	ACCAT	TGCT	AC	CGG	GA	AA	CC		861							
Db	904	GCCAT	TGCC	AC	CGG	GA	AA	CC		924							

Search completed: March 31, 2005, 00:50:27  
Job time : 924 secs

1875

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 31, 2005, 00:32:47 ; Search time 322 Seconds  
(without alignments)  
8852.155 Million cell updates/sec

Title: US-10-049-822A-1  
Perfect score: 1742  
Sequence: 1 cacagagcagaccagtgag.....ggagccctgggggttcctc 1742

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PTCUS COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	741	42.5	926	3	US-08-481-739-1
2	741	42.5	926	5	PCT-US94-07089-6
3	737.8	42.4	926	1	US-08-081-448-5
4	737.8	42.4	926	2	US-08-470-670A-6
5	737.8	42.4	926	3	US-09-167-921-1
6	737.8	42.4	926	3	US-09-271-020-39
7	737.8	42.4	926	3	US-08-323-743-1
8	737.8	42.4	926	3	US-08-461-511A-6
9	737.8	42.4	926	4	US-09-023-655-1430
10	737.8	42.4	926	4	US-09-814-915A-106
11	736.2	42.3	926	3	US-09-271-014A-5
12	627.6	36.0	1236	4	US-09-639-245-1
13	599.2	34.4	1384	4	US-08-899-367-1
14	576.8	33.1	60489	4	US-09-949-016-16287
15	563	32.3	1455	4	US-09-639-245-7
16	399	22.9	737	4	US-09-949-016-4545
17	399	22.9	737	4	US-09-023-655-1458
18	399	22.9	737	5	PCT-US94-07089-8
19	395.8	22.7	737	1	US-08-081-448-7
20	395.8	22.7	737	2	US-08-470-670A-8
21	395.8	22.7	737	3	US-08-461-511A-8
22	395.8	22.7	737	3	US-09-271-014A-7
23	338.4	19.4	1274	2	US-08-470-670A-1
24	338.4	19.4	1274	3	US-08-461-511A-1
25	338.4	19.4	1274	5	PCT-US94-07089-1
26	338.4	19.4	1303	1	US-08-081-448-1
27	338.4	19.4	1303	3	US-09-271-014A-1

28 155.2 8.9 717 2 US-08-465-485A-20 Sequence 20, Appl  
29 155.2 8.9 717 3 US-09-080-285-20 Sequence 20, Appl  
30 155.2 8.9 4825 6 5459251-1 Patent No. 5459251  
31 155.2 8.9 4825 6 5459251-1 Patent No. 5459251  
32 155.2 8.9 5086 2 US-08-465-485A-19 Sequence 19, Appl  
33 155.2 8.9 5086 2 US-08-365-486A-14 Sequence 14, Appl  
34 155.2 8.9 5086 3 US-09-080-285-19 Sequence 19, Appl  
35 155.2 8.9 5086 3 US-08-880-342-14 Sequence 14, Appl  
36 155.2 8.9 5086 3 US-09-724-426-19 Sequence 19, Appl  
37 155.2 8.9 5086 3 US-09-233-527-7 Sequence 7, Appl  
38 155.2 8.9 5086 5 PCT-US93-05651-4 Sequence 4, Appl  
39 155.2 8.9 5086 5 PCT-US93-06251-2 Sequence 2, Appl  
40 155.2 8.9 5094 3 US-09-234-186-7 Sequence 7, Appl  
41 155.2 8.9 5104 6 5506344-1 Patent No. 5506344  
42 155.2 8.9 5104 6 5506344-1 Patent No. 5506344  
43 154.2 8.9 1846 2 US-08-365-486A-16 Sequence 16, Appl  
44 154.2 8.9 1846 3 US-08-880-342-16 Sequence 16, Appl  
45 154.2 8.9 6030 4 US-09-023-655-1015 Sequence 1015, Ap

ALIGNMENTS

RESULT 1  
US-08-481-739-1  
; Sequence 1, Application US/08481739  
; Patent No. 6143291  
; GENERAL INFORMATION:  
; APPLICANT: June, Carl H. and Thompson, Craig B.  
; TITLE OF INVENTION: METHODS FOR ENHANCING T CELL SURVIVAL  
; TITLE OF INVENTION: BY AUGMENTING BCL-XL PROTEIN LEVELS  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 28 State Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/481,739  
; FILING DATE: 07-JUNE-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/435,518  
; FILING DATE: 04-MAY-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: DeConti, Giulio A. (GAD)  
; REGISTRATION NUMBER: 31,503  
; REFERENCE/DOCKET NUMBER: RPI-034CP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617)227-7400  
; TELEFAX: (617)742-4214  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 926 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 135..836  
; US-08-481-739-1

Query Match 42.5%; Score 741; DB 3; Length 926;  
Best Local Similarity 91.3%; Pred. No. 1.2e-218;  
Matches 786; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

```

Qy 1 CACAGAGCAGACCCAGTGTGAGCAGGTGTTTGGACAATGGACTGGTTGAGCCCATCT 60
Db |||||
Qy 64 CAGAGCGAGACTCAGTGTGAGCAGGTGTTTGGACAATGGACTGGTTGAGCCCATCC 123
Db |||||
Qy 61 CTATTATAAATGTCTCAGAGCAACCGGAGCTGGTGTGACTTTCTCTCTACAGC 120
Db |||||
Qy 124 CTATTATAAATGTCTCAGAGCAACCGGAGCTGGTGTGACTTTCTCTCTACAGC 183
Db |||||
Qy 121 TCTCCCAAGAGGATACAGCTGGAGTCACTTTAGCGATGTGAGAGAGAGGAGTGAAG 180
Db |||||
Qy 184 TTTCCCAAGAGGATACAGCTGGAGTCACTTTAGTGATGTGGAAGAGAGGAGTGAAG 243
Db |||||
Qy 181 CCCAGAGAAATCAACAGAGGAGAGACCCAGTGCCCATCAATGGCAACCCATCCT 240
Db |||||
Qy 244 CCCAGAGGAGTCAATCGAGATGGAGACCCAGTGCCCATCAATGGCAACCCATCCT 303
Db |||||
Qy 241 GGCACTGGCGGATAGCCCGCGGTGAATGAGGCCACTGGCCACAGCAGCAGTTTGGATG 300
Db |||||
Qy 304 GGCACTGGCGGATAGCCCGCGGTGAATGAGGCCACTGGCCACAGCAGCAGTTTGGATG 363
Db |||||
Qy 301 CGCGGAGGTAAATCCCATGCGAGCAGTGAAGCAAGCGCTGAGAGAGCTGGCATGAGT 360
Db |||||
Qy 364 CCCGGGAGGTGATCCCCATGCGAGCAAGTAAAGCAAGCGCTGAGGAGGAGCGAGCAGT 423
Db |||||
Qy 361 TTGAATCGCGTACCGGAGCAGTTCAGTGATCTAAATCCAGCTTCATATAACCCAG 420
Db |||||
Qy 424 TTGAATCGCGTACCGGCGGCACTCAGTGACCTGACATCCAGCTCCATATCACCCAG 483
Db |||||
Qy 421 GGACAGCATATCAGAGCTTTGAACAGAGTAGTGAATGAATCTTTTCGGGATGGGTAACT 480
Db |||||
Qy 484 GGACAGCATATCAGAGCTTTGAACAGAGTAGTGAATGAATGAATCTTTTCGGGATGGGTAACT 543
Db |||||
Qy 481 GGGGTGCGATTTGGCCCTTTCTTCCTTTGGCGGGGCACTGTGCGTGAAGAGCGTAGACA 540
Db |||||
Qy 544 GGGGTGCGATTTGGCCCTTTTCTCTTCGGCGGGGCACTGTGCGTGAAGAGCGTAGACA 603
Db |||||
Qy 541 AGGAGATCGAGTATTGTGTGAGTGGATTGCAAGTTGGATGCGCCACCTTACCTGAATGACC 600
Db |||||
Qy 604 AGGAGATCGAGTATTGTGTGAGTGGATTGCAAGTTGGATGCGCCACCTTACCTGAATGACC 663
Db |||||
Qy 601 ACCTAGAGCTTTGGATCCAGGAGAACCGCGGCTGGGACACTTTTGTGGATCTCTACGGGA 660
Db |||||
Qy 664 ACCTAGAGCTTTGGATCCAGGAGAACCGCGGCTGGGATACCTTTTGTGGATCTCTATGGGA 723
Db |||||
Qy 661 ACAATGAGAGCGAGAGCGGAAAGCGGAGAGCGTTTCAACCGCTGGTTCCTGACGG 720
Db |||||
Qy 724 ACAATGAGAGCGAGAGCGGAAAGCGGAGAGCGTTTCAACCGCTGGTTCCTGACGG 783
Db |||||
Qy 721 GCATGACTGTGGCTGGTGTGATTTCTGTGGCTCACTCTTCAGTGGGAAGTGACAGACA 780
Db |||||
Qy 784 GCATGACTGTGGCTGGTGTGATTTCTGTGGCTCACTCTTCAGTGGGAATGACAGACA 843
Db |||||
Qy 781 CTGACCGTCCACTCACTCTCACTCCACCTTCCCGGCGGCGGCAACCACTCTCTTTCAGCC 840
Db |||||
Qy 844 CTGACCTCCACTCTACCTCTCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 861
Db |||||
Qy 841 ACCATTGCTACCGAGAGAAC 861
Db |||||
Qy 904 GCCATTGCCACCGAGAGAAC 924

```

RESULT 2

```

PCT-US94-07089-6
; Sequence 6, Application PC/TUS9407089
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Vertebrate Apoptosis Gene:
; TITLE OF INVENTION: Compositions and Methods
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston

```

```

; STATE: TX
; COUNTRY: United States of America
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/07089
; FILING DATE: CONCURRENTLY FILED
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/081.448
; FILING DATE: 22 JUNE 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: PARKER, David L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: ARCD090
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512-320-7200
; TELEFAX: 713-789-2679
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 926 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 135..836
; PCT-US94-07089-6

```

```

Query Match 42.5%; Score 741; DB 5; Length 926;
Best Local Similarity 91.3%; Pred. No. 1.2e-218;
Matches 786; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

Qy 1 CACAGAGCAGACCCAGTGTGAGCAGGTGTTTGGACAATGGACTGGTTGAGCCCATCT 60
Db |||||
Qy 64 CAGAGCGAGACTCAGTGTGAGCAGGTGTTTGGACAATGGACTGGTTGAGCCCATCC 123
Db |||||
Qy 61 CTATTATAAATGTCTCAGAGCAACCGGAGCTGGTGTGACTTTCTCTCTACAGC 120
Db |||||
Qy 124 CTATTATAAATGTCTCAGAGCAACCGGAGCTGGTGTGACTTTCTCTCTACAGC 183
Db |||||
Qy 121 TCTCCCAAGAGGATACAGCTGGAGTCACTTTAGCGATGTGAGAGAGAGGAGTGAAG 180
Db |||||
Qy 184 TTTCCCAAGAGGATACAGCTGGAGTCACTTTAGTGATGTGGAAGAGAGGAGTGAAG 243
Db |||||
Qy 181 CCCAGAGAAATCAACAGAGGAGAGACCCAGTGCCCATCAATGGCAACCCATCCT 240
Db |||||
Qy 244 CCCAGAGGAGTCAATCGAGATGGAGACCCAGTGCCCATCAATGGCAACCCATCCT 303
Db |||||
Qy 241 GGCACTGGCGGATAGCCCGCGGTGAATGAGGCCACTGGCCACAGCAGCAGTTTGGATG 300
Db |||||
Qy 304 GGCACTGGCGGATAGCCCGCGGTGAATGAGGCCACTGGCCACAGCAGCAGTTTGGATG 363
Db |||||
Qy 301 CGCGGAGGTAAATCCCATGCGAGCAGTGAAGCAAGCGCTGAGAGAGCTGGCATGAGT 360
Db |||||
Qy 364 CCCGGGAGGTGATCCCCATGCGAGCAAGTAAAGCAAGCGCTGAGGAGGAGCGAGCAGT 423
Db |||||
Qy 361 TTGAATCGCGTACCGGAGCAGTTCAGTGATCTAAATCCAGCTTCATATAACCCAG 420
Db |||||
Qy 424 TTGAATCGCGTACCGGCGGCACTCAGTGACCTGACATCCAGCTCCATATCACCCAG 483
Db |||||
Qy 421 GGACAGCATATCAGAGCTTTGAACAGAGTAGTGAATGAATCTTTTCGGGATGGGTAACT 480
Db |||||
Qy 484 GGACAGCATATCAGAGCTTTGAACAGAGTAGTGAATGAATGAATCTTTTCGGGATGGGTAACT 543
Db |||||
Qy 481 GGGGTGCGATTTGGCCCTTTCTTCCTTTGGCGGGGCACTGTGCGTGAAGAGCGTAGACA 540
Db |||||
Qy 544 GGGGTGCGATTTGGCCCTTTTCTCTTCGGCGGGGCACTGTGCGTGAAGAGCGTAGACA 603
Db |||||

```





CORRESPONDENCE ADDRESS:			
ADDRESSEE: Arnold, White & Durkee			
STREET: P.O. Box 4433			
CITY: Houston			
STATE: Texas			
COUNTRY: United States of America			
ZIP: 77210			
COMPUTER READABLE FORM:			
MEDIUM TYPE: Floppy disk			
COMPUTER: IBM PC compatible			
OPERATING SYSTEM: PC-DOS/MS-DOS			
SOFTWARE: PatentIn Release #1.0, Version #1.30			
CURRENT APPLICATION DATA:			
APPLICANT NUMBER: US/08/470,670A			
FILING DATE:			
CLASSIFICATION: 514			
PRIOR APPLICATION DATA:			
APPLICATION NUMBER: US 08/081,448			
FILING DATE: 22-JUN-1993			
ATTORNEY/AGENT INFORMATION:			
NAME: Highlander, Steven L.			
REGISTRATION NUMBER: 37,642			
REFERENCE/DOCKET NUMBER: ARCD:090--1			
TELECOMMUNICATION INFORMATION:			
TELEPHONE: (512) 418-3000			
TELEFAX: (512) 474-7577			
INFORMATION FOR SEQ ID NO: 6:			
SEQUENCE CHARACTERISTICS:			
LENGTH: 926 base pairs			
TYPE: nucleic acid			
STRANDEDNESS: single			
TOPOLOGY: linear			
FEATURE:			
NAME/KEY: CDS			
LOCATION: 135..836			
US-08-470-670A-6			
Query Match 42.4%; Score 737.8; DB 2; Length 926;			
Best Local Similarity 91.1%; Pred. No. 1.2e-217;			
Matches 784; Conservative 0; Mismatches 77; Indels 0; Gaps 0;			
Qy	1	CACAGACGACCCTGAGTGGACGAGGTGTTTGGAACAATGCAGTGGTGGCCCATCT	60
Db	64	CAGACGAGACTCAGTAGTGAGCAGGTGTTTGGAACAATGCAGTGGTGGCCCATCT	123
Qy	61	CTATTATAAAATGTCTCAGAGCAACCGGAGCTGGTGGTTCACCTTCTCTCTACAGC	120
Db	124	CTATTATAAAATGTCTCAGAGCAACCGGAGCTGGTGGTTCACCTTCTCTCTACAGC	183
Qy	121	TCTCCAGAAGAGATACAGTGGAGTCACTTTAGCGATGTCGAAGAACAGGACTGAAG	180
Db	184	TTTTCCAGAAAAGATACAGTGGAGTCACTTTAGCGATGTCGAAGAACAGGACTGAAG	243
Qy	181	CCCCAGAAGAACTGAACAGAAAAGGAGACCCCGAGTGCACATCAATGCCAACCCATCT	240
Db	244	CCCCAGAAGAGTGAATCGAGATGGAGACCCCGAGTGCACATCAATGCCAACCCATCT	303
Qy	241	GGCACCTGGCGGATAGCCCCCGGGTGAATGGAGCCACTGGCCACAGCAGCAGTTGGATG	300
Db	304	GGCACCTGGCAGACACCCCGGGTGAATGGAGCCACTGGCCACAGCAGCAGTTGGATG	363
Qy	301	CGCGGAGGTAATCCCCATGGCAGTAGTGAAGCAACCGCTGAGAGAGCTGGCGATGAGT	360
Db	364	CGCGGAGGTAATCCCCATGGCAGTAGTGAAGCAACCGCTGAGAGAGCTGGCGAGT	423
Qy	361	TTGAACTGCGGTACCGGAGAGCAATTCAGTGATCTAACATCCCAGCTTCATATAACCCAG	420
Db	424	TTGAACTGCGGTACCGGAGAGCAATTCAGTGATCTAACATCCCAGCTTCATATAACCCAG	483
Qy	421	GGAAGCATATCAGAGCTTTTGAACAGGTAGTGAATGAATCTTTTCGGATGGGGTAAACT	480
Db	484	GGAAGCATATCAGAGCTTTTGAACAGGTAGTGAATGAATCTTTTCGGATGGGGTAAACT	543



```

; FILE REFERENCE: ISPH-0368
; CURRENT APPLICATION NUMBER: US/09/323,743
; CURRENT FILING DATE: 1999-06-01
; EARLIER APPLICATION NUMBER: 09/277,020
; EARLIER FILING DATE: 1998-03-26
; EARLIER APPLICATION NUMBER: 09/167,921
; EARLIER FILING DATE: 1998-10-07
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 926
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (135)..(836)
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: L20121 Genbank
; DATABASE ENTRY DATE: 1994-07-26
; US-09-323-743-1

```

Query Match	42.4%	Score 737.8	DB 3	Length 926
Best Local Similarity	91.1%	Pred. No. 1.2e-217		
Matches 784	Conservative 0	Mismatches 77	Indels 0	Gaps 0

Qy	1	CACAGACGACCCAGTGAGTGACGAGGTGTTTTGGACAAATGGACTGGTTGAGCCCACTCT	60
Db	64	CAGAGACGACACTCAGTGAGTGACGAGGTGTTTTGGACAAATGGACTGGTTGAGCCCACTCC	123
Qy	61	CTATTATAAAAATGCTCTCAGAGCAACCGGGAGCTGGTGGTTGACTTTCTCTCTCAACAAGC	120
Db	124	CTATTATAAAAATGCTCTCAGAGCAACCGGGAGCTGGTGGTTGACTTTCTCTCTCAACAAGC	183
Qy	121	TCTCCCAAGAAAGGATACAGCTGGAGTCAGTTTACGATGTGCAAGAGAAACAGGACTGAAG	18
Db	184	TTTCCCAAGAAAGGATACAGCTGGAGTCAGTTTACGATGTGCAAGAGAAACAGGACTGAGG	243
Qy	181	CCCAGAGAAACTGAACCAAGAGGGAGACCCCACTGTCCTCAATGGCAACCCATCCT	240
Db	244	CCCAGAAAGGACTGAATCGAGATGAGACCCCACTGTCCTCAATGGCAACCCATCCT	303
Qy	241	GGCACTTGGGGATAGCGCCCGCGTGAATGAGGCCACTGGGCCACAGACAGCAGTTTGGATG	300
Db	304	GGCACTTGGCAGACAGCCCCGCGTGAATGGAGCCACTGGGCCACAGCAGCAGTTTGGATG	363
Qy	301	CGCGGAGGTAAATCCCATGCGAGTGGAAGCAAGCGCTGAGAGAGGCTGGCGATGAGT	360
Db	364	CCCGGAGGTGATCCCATGCGCAGCTAAAGCAAGCGCTGAGGAGGCGAGGCGACGAGT	423
Qy	361	TTGNACTGCGGTACCGGAGGAGCTTCAGTGATCTAAACATCCAGCTTCATATACCCCGAG	420
Db	424	TTGNACTGCGGTACCGGCGGCACTTCAGTGACCTTGACATCCAGCTCCACATCAACCCCGAG	483
Qy	421	GGACAGCATATCAGAGCTTTGAAACAGGTAGTGAATGAACTCTTTCCGGGATGGGCTAAACT	480
Db	484	GGACAGCATATCAGAGCTTTGAAACAGGTAGTGAATGAACTCTTTCCGGGATGGGCTAAACT	543
Qy	481	GGGGTCGATGTGGCTCTTCTCTCTTTGGCGGGGCACTGTGCGTGGAAAGCGTGTAGACA	540
Db	544	GGGGTCGATGTGGCTCTTCTCTCTTTGGCGGGGCACTGTGCGTGGAAAGCGTGTAGACA	603
Qy	541	AGGAGATGCAGGTATTTGGTGAGTCGGATTTGCAAGTTGGATGGCCCACTTACCTGATGACC	600
Db	604	AGGAGATGCAGGTATTTGGTGAGTCGGATTCGGCAGCTTGGATGGCCCACTTACCTGATGACC	663
Qy	601	ACCTAGAGCCTTTGGATCCAGAGAAACGGCGGCTGGGACATTTTGTGTGATCTCTTACGGGA	660
Db	664	ACCTAGAGCCTTTGGATCCAGAGAAACGGCGGCTGGGATCTTTTGTGTGAACTCTTATGGGA	723
Qy	661	ACAATGACGACCGCAGAGCGCGGAAAGGCCACAGGAGCGTTTCAACCGCTGGTTCTCTGACGG	720
Db	724	ACAATGACGACCGCAGAGCGCGGAAAGGCCACAGGAAACGTTTCAACCGCTGGTTCTCTGACGG	783

Qy	721	GCATGACTGTGGCTGGTGTAGTTCTGCTGGGCTCACTCTTCAGTCGGAAGTGA	CCAGACA	780
Db	784	GCATGACTGTGGCGGGCGTGTCTGCTGGGCTCACTCTTCAGTCGGAATGACCAGACA	843	
Qy	781	CTGACCTGCTCACTCACTCTCACTCTCCACCTTCCACCTTCCCGCCACCACTCTCTCTCTTCAGCC	840	
Db	844	CTGACCATCTCACTTACCTTCCACCTCCCTCTCTGCTTCCACCATCTCTCGTCCAGCC	903	
Qy	841	ACCATTGCTACCGAGAACC	861	
Db	904	GCCATTGCCACGAGAAC	924	

RESULT 8  
 US-08-461-511A-6  
 ; Sequence 6, Application US/08461511A  
 ; Patent No. 6303331  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Thompson, Craig B.B.  
 ; Boise, Lawrence H.  
 ; TITLE OF INVENTION: VERTEBRATE APOPTOSIS GENE: COMPOSITIONS  
 ; AND METHODS  
 ; NUMBER OF SEQUENCES: 18  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Arnold, White & Durkee  
 ; STREET: P.O. Box 4433  
 ; CITY: Houston  
 ; STATE: Texas  
 ; COUNTRY: United States of America  
 ; ZIP: 77210  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/461,511A  
 ; FILING DATE: 05-Jun-1995  
 ; CLASSIFICATION: UNKNOWN  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Highlander, Steven L.  
 ; REGISTRATION NUMBER: 37,642  
 ; REFERENCE/DOCKET NUMBER: ARCD:179  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (512) 418-3000  
 ; TELEFAX: (512) 474-7577  
 ; INFORMATION FOR SEQ ID NO: 6:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 926 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 135..836  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 6:  
 US-08-461-511A-6

	Query Match	42.4%	Score 737.8	DB 3	Length 926
	Best Local Similarity	91.1%	Pred. No. 1.2e-217		
	Matches 784	Conservative 0	Mismatches 77	Indels 0	Gaps 0
Qy	1	CACAGACGAGCCACGTGAGTGACGAGGTGTTTTGGACAAATGGACTGCTTGAGCCCATCT	60		
Db	64	CAGAGACGAGACTCAGTGAGTGACGAGGTGTTTTGGACAAATGGACTGCTTGAGCCCATCC	123		
Qy	61	CTATTATAAAAAATGCTCTCAGAGCAACCGGGAGCTGGTGGTTGACTTTCTCTCTACAAGC	120		
Db	124	CTATTATAAAAAATGCTCTCAGAGCAACCGGGAGCTGGTGGTTGACTTTCTCTCTACAAGC	183		
Qy	121	TCTCCCGAAAGGATACAGCTGGAGTTCAGTTTAGCGATGTCGAGAGAAACAGGACTGTAAG	180		
Db	184	TTTCCCGAAAGGATACAGCTGGAGTTCAGTTTAGTGTGGAAGAGAAACAGGACTGTAAG	243		







```

; APPLICATION NUMBER: US/08/899,367
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: AMY E. Mandragouras
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: DFN-019
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1384 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 378..1085
US-08-899-367-1

```

Query Match	34.4%;	Score 599.2;	DB 4;	Length 1384;
Best Local Similarity	96.4%;	Pred. No. 1.2e-174;		
Matches 613;	Conservative 0;	Mismatches 23;	Indels 0;	Gaps 0;
QY	1	CACAGACAGACCCAGTGGAGTGACGAGTGTCTTTGGACAATGGACTGGTTGAGCCCATCT	60	
DB	307	CACAGACAGACCCAGTAAAGTGACGAGGTGTTTGGACATGGACTGGTTGAGCCCATCT	366	
QY	61	CTATTATAAAATGCTTCAGAGCAACCGGGAGCTGGTGGTTGAATTCTCTCTTACAAAGC	120	
DB	367	CTATTATAAAATGCTTCAGAGCAACCGGGAGCTGGTGGTGCATCTTCTCTCTTACAAAGC	426	
QY	121	TCTCCCAAGAAAGGATACAGCTGGAGTTCAGTTTACGATGTGCAAGAGAAACAGGACTGAAG	180	
DB	427	TTTCCCAAGAAAGGATACAGCTGGAGTTCAGTTTACGATGTGCAAGAGAAATAGGACTGAGG	486	
QY	181	CCCAGAGAAACTGAAACCGAAAGGGAGACCCCGAGTGCATCAATGGCAACCATCTCT	240	
DB	487	CCCAGAGAAACTGAAAGCAGAGAGGAGACCCCGAGTGCATCAATGGCAACCATCTCT	546	
QY	241	GGCACCTGGCGGATAGCCCGCGTGAATGGAGCCACTGGGCCACAGCAGCAGGTTTGGATG	300	
DB	547	GGCACCTGGCGGATAGCCCGCGCGTGAATGGAGCCACTGGGCCACAGCAGCAGGTTTGGATG	606	
QY	301	CGCGGAGGTAAATCCCATGGCAGCAGTGAAGCAAGCGCTCAGAGAGCGCTGCGGATGAGT	360	
DB	607	CGCGGAGGTGATTTCCCATGGCAGCAGTGAAGCAAGCGCTCAGAGAGCGCAGCGGATGAGT	666	
QY	361	TTGAACTGCGGTACCGGAGAGCATTCAGTGATCTAAACATCCAGGTTCAATAACCCGAG	420	
DB	667	TTGAACTGCGGTACCGGAGAGCGTTCAAGTGAATCAACATCCAGCTTCACATAACCCGAG	726	
QY	421	GGACAGCATATCAGAGCTTTGAAACAGGTAAGTAATGAACTCTTTGGGATGGGCTAAACT	480	
DB	727	GGACCGGTATCAGAGCTTTGAGCAGGTAGTGAAATGAATGAATCTTTTCGGGATGGGTTAACT	786	
QY	481	GGGTCGGAATGTGGCTTTCTTCTTTTGGCGGGGCACTGTGGTGGAAGCGGTAGACA	540	
DB	787	GGGTCGCAATCGTGGCCCTTTTCTCTTTTGGCGGGGCACTGTGGTGGAAGCGGTAGACA	846	
QY	541	AGGAGATGCAGGTATTTGGTGAGTCCGATTGCAAGTTGGATGGCCACCTPACTCGAATGACC	600	
DB	847	AGGAGATGCAGGTATTTGGTGAGTCCGATTGCAAGTTGGATGGCCACCTPACTCGAATGACC	906	
QY	601	ACTTAGAGCCTTGGATCCAGAGAACCGCGGCTGGG	636	
DB	907	ACTTAGAGCCTTGGATCCAGAGAACCGCGGCTGGG	942	

RESULT 14

```

US-09-949-016-16287
; Sequence 16287, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 16287
; LENGTH: 60489
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16287

Query Match          33.1%; Score 576.8; DB 4; Length 60489;
Best Local Similarity 94.2%; Pred. No. 1.1e-166;
Matches 599; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

Qy 1 CACAGACGACCCAGTGTGAGTGACGAGTGTTTTGGACAATGGACTGTGTGAGCCCATCT 60
Db 2064 CAGAGACGAGACTCAGTGTGAGTGAGCAGTGTTTTGGACAATGGACTGTGTGAGCCCATCC 2123

Qy 61 CTATTATAAAATGCTCTCAGAGCAACCGGAGCTGGTGTGACTTTCTCTCTACACAGC 120
Db 2124 CTATTATAAAATGCTCTCAGAGCAACCGGAGCTGGTGTGACTTTCTCTCTACACAGC 2183

Qy 121 TCTCCGAGAAGGATACAGCTGGAGTCAGTGTTCAGTGTGTCGAAGAGAACAGGACTGAAG 180
Db 2184 TTTCCGAGAAGGATACAGCTGGAGTCAGTGTTCAGTGTGTCGAAGAGAACAGGACTGAGG 2243

Qy 181 CCCCAGAAAGAACTGAACACGAAAGGAGGACCCCGAGTGCCTCAATGGCAACCCATCCT 240
Db 2244 CCCCAGAAAGGAACTGAATCGGAGATGAGACCCCGAGTGCCTCAATGGCAACCCATCCT 2303

Qy 241 GGCACTGGCGGATAGCCCGCGTGAATGGAGCCACTGGGCCACAGCAGCAGTTTGGATG 300
Db 2304 GGCACCTGGCAGACAGCCCGCGTGAATGGAGCCACTGGGCCACAGCAGCAGTTTGGATG 2363

Qy 301 CGCGGAGGTAATCCCATGGCAGCAGTGAAGCAAGCGCTGAGAGAGGCTGGCGATGAGT 360
Db 2364 CCCGGGAGTGAATCCCATGGCAGCAGTGAAGCAAGCGCTGAGGAGGCGAGCGACGAGT 2423

Qy 361 TTGACTCGCGTACCGGAGGAGCATTCAGTGAATCTAAACATCCCGCTTCATATAACCCAG 420
Db 2424 TTGAATCGCGGTACCGGCGGCATTCAGTGAATCTAAACATCCCGCTTCATATAACCCAG 2483

Qy 421 GGACAGCATATCAGAGCTTTGAAACAGGTAGTGAATGAACCTTTTCGGGATGGGGTAAACT 480
Db 2484 GGACAGCATATCAGAGCTTTGAAACAGGTAGTGAATGAACCTTTTCGGGATGGGGTAAACT 2543

Qy 481 GGGGTGCGATGTGGCCCTTCTTCTCTTTGCGGGGCACTGTGGTGGAAAGCGTAGACA 540
Db 2544 GGGGTGCGATGTGGCCCTTCTTCTCTTTGCGGGGCACTGTGGTGGAAAGCGTAGACA 2603

Qy 541 AGGAGATGCAAGTATTGCTGAGTCGGAATTCGAAGTTGGATGCCACCTTACCTGAATGACC 600
Db 2604 AGGAGATGCAAGTATTGCTGAGTCGGAATTCGAAGTTGGATGCCACCTTACCTGAATGACC 2663

Qy 601 ACCTAGAGCCTTGGATCCAGAGAACCGCGCTCGG 636
Db 2664 ACCTAGAGCCTTGGATCCAGAGAACCGCGCTCGG 2699

```



## RESULT 15

US-09-639-245-7  
; Sequence 7, Application US/09639245  
; Patent No. 6737511  
; GENERAL INFORMATION:  
; APPLICANT: Youle et al.  
; TITLE OF INVENTION: RECEPTOR-MEDIATED UPTAKE OF AN EXTRACELLULAR BCL-XL  
; FILE REFERENCE: 4239-55417  
; CURRENT APPLICATION NUMBER: US/09/639,245  
; CURRENT FILING DATE: 2000-08-15  
; PRIOR APPLICATION NUMBER: 60/149,220  
; PRIOR FILING DATE: 1999-08-16  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 7  
; LENGTH: 1455  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: genetic fusion  
; NAME/KEY: CDS  
; LOCATION: (1)..(1455)  
US-09-639-245-7

Query Match 32.3%; Score 563; DB 4; Length 1455;

Best Local Similarity 92.9%; Pred. NO. 2e-163;

Matches 590; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

Qy	63	ATTATAAATGCTCAGAGCAACGGGAGCTGGTGTGACTTTCTCTCTACAGCTC	122
Db	820	AATCTATCCATGCTCAGAGCAACGGGAGCTGGTGTGACTTTCTCTCTACAGCTT	879
Qy	123	TCCAGAAAGGATACAGCTGGAGTCAGTTTAGCGATGTCGAAGAGAACAGGACTGAAGCC	182
Db	880	TCCAGAAAGGATACAGCTGGAGTCAGTTTAGCTAGTGTGAAGAGAACAGGACTGAAGCC	939
Qy	183	CCAGAAGAACTGAACCCAGAAAGGAGACCCCACTGTCATCAATGGCAACCCATCTCGG	242
Db	940	CCAGAAGGACTGAATCGGAGATGGAGACCCCACTGTCATCAATGGCAACCCATCTCGG	999
Qy	243	CACCTGGGGATAGCCCGCGGTGAATGGAGCCACTGGCCACAGCAGAGTTTGGATGCG	302
Db	1000	CACCTGGCAGACAGCCCGCGGTGAATGGAGCCACTGGCCACAGCAGAGTTTGGATGCG	1059
Qy	303	CGGAGGTATCCCATGGCAGCAGTGAAGCAAGCGCTGAGAGAGGCTGGCGATGAGTTT	362
Db	1060	CGGAGGTATCCCATGGCAGCAGTGAAGCAAGCGCTGAGAGAGGCTGGCGATGAGTTT	1119
Qy	363	GAACTGGCGTACCGAGAGCAATTCAGTGATCTAACATCCAGCTTTCATATAACCCAGGG	422
Db	1120	GAACTGGCGTACCGAGAGCAATTCAGTGATCTAACATCCAGCTTTCATATAACCCAGGG	1179
Qy	423	ACAGCATATCAGAGTTTGAACAGTAGTGAATGAATCTTTTCGGGATGGGGTAACTGG	482
Db	1180	ACAGCATATCAGAGTTTGAACAGTAGTGAATGAATCTTTTCGGGATGGGGTAACTGG	1239
Qy	483	GGTCGATTTGGGCTTCTCTCTTTGGCGGGCAGTGGGTGGAAAGGCTAGACAAAG	542
Db	1240	GGTCGATTTGGGCTTCTCTCTTTGGCGGGCAGTGGGTGGAAAGGCTAGACAAAG	1299
Qy	543	GAGATGCAGGTATTTGGTGGATTCGAGTTCGAGTTGGATGGCCACCTCAATGAATGACCA	602
Db	1300	GAGATGCAGGTATTTGGTGGATTCGAGTTCGAGTTGGATGGCCACCTCAATGAATGACCA	1359
Qy	603	CTAGAGCCTTGGATCCAGAGAAACGGCGGCTGGGACACTTTTGTGGATCTCTACGGGAAC	662
Db	1360	CTAGAGCCTTGGATCCAGAGAAACGGCGGCTGGGACACTTTTGTGGATCTCTATGGGAAC	1419
Qy	663	AATGCAGCCGAGAGCCGAAAGGCCAGAGCG	697
Db	1420	AATGCAGCCGAGAGCCGAAAGGCCAGAGCG	1454

Search completed: March 31, 2005, 04:26:50  
Job time : 325 secs

THE GOVERNMENT

	Query Match	67.08;	Score 1168;	DB 18;	Length 1466;
	Best Local Similarity	91.09;	Pred. No. 0;		
	Matches 1331;	Conservative	0;	Mismatches 95;	Indels 36; Gaps
Qy	1	CACAGACGACGCCAGTGGAGTGGACAGGTGTTTTGGCAATGGACTGGTTGAGGCCCATCT	60		
Db	32	CACAGACGACGCCAGTAACTGGACAGGTGTTTTGGCAATGGACTGGTTGAGGCCCATCT	91		

Qy 61 CTATTATAAAATGTCTCAGAGCAACCGGAGCTGGTGGTTGACTTTCTCTCTCTACAAGC 120  
Db |||||  
Qy 92 CTATTATAAAATGTCTCAGAGCAACCGGAGCTGGTGGTTGACTTTCTCTCTCTACAAGC 151  
Db |||||  
Qy 121 TCTCCGAAAGGATACAGCTGGAGTCAAGTTAGCGATGTGCAAGAGAACAGGACTGAAG 180  
Db |||||  
Qy 152 TTTCCGAAAGGATACAGCTGGAGTCAAGTTAGTGATGTGCAAGAGAAATAGGACTGAGG 211  
Db |||||  
Qy 181 CCCAGAGAAACTGAAACAGAAAGGAGACCCCGAGTGGCCATCAATGGCAACCATCTCT 240  
Db |||||  
Qy 212 CCCAGAGAAACTGAAACAGAGAGGAGACCCCGAGTGGCCATCAATGGCAACCATCTCT 271  
Db |||||  
Qy 241 GGCACCTGGCGATAGCCCGGCTGAATGAGCCACTTGGCCACAGCAGCAGTGTGGATG 300  
Db |||||  
Qy 272 GGCACCTGGCGATAGCCCGGCTGAATGAGCCACTTGGCCACAGCAGCAGTGTGGATG 331  
Db |||||  
Qy 301 CGCGGAGGTTAATCCCATGTCAGCAGTGAAGCAAGCGCTGAGAGAGCTGGCGATGAGT 360  
Db |||||  
Qy 332 CGCGGAGGTTAATCCCATGTCAGCAGTGAAGCAAGCGCTGAGAGAGCGAGCGATGAGT 391  
Db |||||  
Qy 361 TTGAACCTGGGTACCGGAGAGCATTCAGTGATCTAAACATCCAGCTTCATATACCCGAG 420  
Db |||||  
Qy 392 TTGAACCTGGGTACCGGAGAGCGTTCACTGATCTAAACATCCAGCTTCACATACCCGAG 451  
Db |||||  
Qy 421 GGACAGCATATCAGAGCTTTGAAACAGGTAGTGAATGAATCTTTTCGGGATGGGTAAACT 480  
Db |||||  
Qy 452 GGACCGGTTATCAGAGCTTTGACGAGGTAGTGAATGAATCTTTTCGGGATGGGTAAACT 511  
Db |||||  
Qy 481 GGGGTGCGATTTGTCCTTTCTCTTTGCGGGGCACTGTGCGTGGAAAGCGTAGACA 540  
Db |||||  
Qy 512 GGGGTGCGATCGTGGCCCTTTCTCTTTGCGGGGCACTGTGCGTGGAAAGCGTAGACA 571  
Db |||||  
Qy 541 AGGAGATCCAGGTATGTGTGAGTGGATGCAAGTTGGATGGCCACCTACCTGAATGACC 600  
Db |||||  
Qy 572 AGGAGATCCAGGTATGTGTGAGTGGATGCAAGTTGGATGGCCACCTATCTGAATGACC 631  
Db |||||  
Qy 601 ACCTAGAGCTTTGGATCCAGAGAACCGCGCTGGGACACTTTTGTGATCTCTACGCGA 660  
Db |||||  
Qy 632 ACCTAGAGCTTTGGATCCAGAGAACCGCGCTGGGACACTTTTGTGATCTCTACGCGA 691  
Db |||||  
Qy 661 ACAATGACGACGCGAGAGCCGGAAGCCAGGAGCGTTTCAACCGCTGGTTCTCTGACGG 720  
Db |||||  
Qy 692 ACAATGACGACGCGAGAGCCGGAAGCAAGGAGGCTTCAACCGCTGGTTCTCTGACGG 751  
Db |||||  
Qy 721 GCATGACTGTGGTGTAGTTCTGTCTGGGCTCACTCTTCACTGCGGAAGTGACAGACA 780  
Db |||||  
Qy 752 GCATGACTGTGGTGTAGTTCTGTCTGGGCTCACTCTTCACTGCGGAAGTGACAGACA 811  
Db |||||  
Qy 781 CTGACCGTCCACTCACTCTCACTCCCACTTGGCCCACTTGGCCCACTCTCTCTTCAGCC 840  
Db |||||  
Qy 812 CTGACCGTCCACTCACTCTCACTCCCACTTGGCCCACTTGGCCCACTCTCTCTTCAGCC 871  
Db |||||  
Qy 841 ACCATTGTACAGGAGAACCACTACATGCAACTCACGCCCTTCCCTATTATAGGGTT 900  
Db |||||  
Qy 872 ACCATTGTACAGGAGAACCACTACATGCAACTCACGCCCTTGGCCCTATCACAGGTT 931  
Db |||||  
Qy 901 GGGCTTACAGGAGTCCCTGAGTTAGCTTTCTAGAACTCTACACGCTTCTGTGAAAGC 960  
Db |||||  
Qy 932 GGGCTTACAGGAGTCCCTGCAATTAGCTTTCTAGGATCTACACACTCTCTGTGAAAGC 991  
Db |||||  
Qy 961 CACCTTCCCGCCACATCTCAGTTCCCTTGGCTCAAAACTCACAGGTTTTCCTTCAGAT 1020  
Db |||||  
Qy 992 CACCTTCCCGCCACATCTCAGTTCTCTTGGCTCAAAACTCACAGGTTTTCCTTCAGAT 1051  
Db |||||  
Qy 1021 CAGCTCTTGGAGGCTGGCAGAGTGGGAAGGGGTGTGTAGAGGAGAGAGCCCTGCCT 1080  
Db |||||  
Qy 1052 CAGCTCTTGGAGGCTGGCAGGAAACAGAGAGGGGTGTG--TAGGGAGAGAGCTGTCT 1108  
Db |||||  
Qy 1081 TGTGGTGGGACCTGTATTAACCTTGAGCC--TCTCGGAATGCTTTTCTGGCAGGAGCT 1139  
Db |||||  
Qy 1109 TGTGGTGGGACCTTAATTAACCTTGAGCCCTTCTCGGAATGCTTTTCCAGCAGGAGCT 1168  
Db |||||  
Qy 1140 GGAGAGCTCTCTAAACCTCTTCCCGCCAGAGACTAGATTGGCTTGGTTTGTGTGTGG 1199

Db 1169 -----CTTCCCTGAAGGAGACTAGCTTGGCTTGGTTTGTATGTGTGG 1213  
Qy 1200 CTTCAGAAATGATTCATTTCCCATTTCTGTGTCTCCCTGGGGGGCTCTCTCTTCCCATC 1259  
Db 1214 CTTCAGACGTGATC--ATTTTCCCATTCGCTGTGTCTCCCTGGG-----TCTCCCTCCTC 1263  
Qy 1260 TCCACCCCGCCCGCCCGCAGAGCCATTGAGTGAGTGTCTTTAGCCCTTTTGACTAACTAA 1319  
Db 1264 CCAATCTCCACCCCGCCCGCAGAGCCATTGAGTGAGTGTCTTTAGCCCTTTTGACTAACTAA 1323  
Qy 1320 AAATTCAGAGCTGTCTTGGGATAAAGGAGCAAGG-----ACCTCTCTCCCGCCTCTGGCCT 1373  
Db 1324 AAATTCAGAGCTGTCTTGGGATAAAGGAGCAAGGACCAAGACCTCTCTCCCGCCTCTGGCCT 1383  
Qy 1374 GGCC-AAGCCCGCCCACTCTCTGTCTGAATGTCTCTGAGGCTCTGTGCTAGAGTCCAGCC 1432  
Db 1384 GGCCAAAACCCCACTTGTGTCTGAATGTCTCTGAGGCTCTGTGCTAGAGTCTAGCC 1443  
Qy 1433 CCACCCAGGAGGAGGACGGAG 1454  
Db 1444 TCACCCAGGAGGAGGAGCTTAG 1465

RESULT 2

US-10-402-017-3  
; Sequence 3, Application US/10402017  
; Publication No. US20030219871A1  
; GENERAL INFORMATION:  
; APPLICANT: Barbara ENENKEL, Heiko MEENTS and Martin FUSSENEGGER  
; TITLE OF INVENTION: Host cells having improved survival properties and methods to gen  
; FILE REFERENCE: Case 1/1314  
; CURRENT APPLICATION NUMBER: US/10402,017  
; CURRENT FILING DATE: 2003-03-28  
; PRIOR APPLICATION NUMBER: US 60/369,307  
; PRIOR APPLICATION NUMBER: April 2, 2002  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 863  
; TYPE: DNA  
; ORGANISM: Cricetulus griseus  
US-10-402-017-3

Query Match 45.2%; Score 787.8; DB 17; Length 863;  
Best Local Similarity 94.6%; Pred. No. 3.5e-237;  
Matches 816; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

Qy 3 CAGAGCAGACCCAGTGAAGTGAAGAGGCTGTTTGGACAATGGACTGGTTGAGCCCTCTCT 62  
Db 1 CAGAGCAGACCCAGTGAAGTGAAGAGGCTGTTTGGACAATGGACTGGTTGAGCCCTCTGT 60  
Qy 63 ATTATAAAATGTCTCAGAGCAACCGGAGCTGGTGGTTGACTTTCTCTCTACAAGCTC 122  
Db 61 ATTATAAAATGTCTCAGAGCAACCGGAGCTAGTGGTTGACTTTCTCTCTACAAGCTC 120  
Qy 123 TCCAGAAAGGATACAGCTGGAGTCAAGTTAGCGATGTTCGAAGAGAAACAGACTGGAAGCC 182  
Db 121 TCCAGAAAGGATACAGCTGGAGTCAAGTTAGTGTATGTTCGAAGAGAAACAGACTGAGGCC 180  
Qy 183 CCAGAAAGAACTGACAGAAAGGAGACCCCGAGTGGCCATCAATGGCAAGCCATCCCTGG 242  
Db 181 CCAGAAAGAACTGAAATCAGAGAGGAGACCCCGAGTGGCCATCAATGGCAAGCCATCCCTGG 240  
Qy 243 CACCTGGCGATAGCCCGCGGTGAATGGAGCCCACTGGCCACAGCAGCAGTGTGGATGCG 302  
Db 241 CACCTGGCGACAGCCCGCGGTGAATGGAGCCCACTGGCCACAGCAGCAGTGTGGATGCA 300  
Qy 303 CGGAGGTAAATCCCATGGCAGCAGTGAAGCAAGCGCTGAGAGAGGCTGGCGATGAGTTT 362  
Db 301 CGGAGGTGATCCCATGGCAGCAGTGAAGCAAGCGCTGAGAGAGGCGCGGATGAGTTT 360



Query Match	42.4%	Score 737.8	DB 9	Length 926
Best Local Similarity	91.1%	Pred. No. 2e-221		
Matches 784	Conservative 0	Mismatches 77	Indels 0	Gaps 0

Qy	1	CACAGACGACCCAGTGTGACGAGTGTGTTGGCAATGGACTGGTTGAGCCCATCT	60
Db			
Qy	64	CAGAGCGAGACTCAGTGTGACGAGTGTGTTGGCAATGGACTGGTTGAGCCCATCC	123
Db			
Qy	61	CTATTATAAAAATCTCTCAGAGCAACCGGAGCTGGTGGTTGACTTTCTCTCCTACAAGC	120
Db			
Qy	124	CTATTATAAAAATGTCTCAGAGCAACCGGAGCTGGTGGTTGACTTTCTCTCTACAAGC	183
Db			
Qy	121	TCTCCCAAGAGATACAGCTGGAGTCAGTTTAGCGATGTGGAAGACAAGCACTGAAG	180
Db			
Qy	184	TTTCCCAAGAGATACAGCTGGAGTCAGTTTAGTGATGTGGAAGACAAGCACTGAGG	243
Db			
Qy	181	CCCCAAGAGAACTGAACCAAGGAAGGAGACCCCGAGTGCCTCAATGGCAACCCATCCT	240
Db			
Qy	244	CCCCAAGAGGACTGAATCGAGATGGAGACCCCGTGCCTCAATGGCAACCCATCCT	303
Db			
Qy	241	GGCACTTGGCGGATAGCCCGGTGAATGGAGCCACTGGGCCACAGCAGCAGTTTGGATG	300
Db			
Qy	304	GGCACTTGGCAGACAGCCCGGTGAATGGAGCCACTGGCCAAGCAGCAGTTTGGATG	363
Db			
Qy	301	CGCGGAGGTAATCCCATGGCAGCTGAAGCAAGCGCTGAGAGAGGCTGGCGATGAGT	360
Db			
Qy	364	CCCGGAGGTGATCCCCATGSCACAGTAAAGCAAGCGCTGAGGAGGCGAGCGCACGAGT	423
Db			
Qy	361	TTGAACTGGCGGTACCGAGAGCATTCAGTGAATCTAAACATCCAGCTTCATATAACCCAG	420
Db			
Qy	424	TTGAACTTGCGGTACCGCGGGCACTTCAGTGACCTTGACATCCAGCTCCACATCACCCCG	483
Db			
Qy	421	GGCAGCATATCAGAGCTTTTGAACAGGTAGTGAATGAATCTTTCCGGATGGGTAAACT	480
Db			
Qy	484	GGCAGCATATCAGAGCTTTTGAACAGGTAGTGAATGAATCTTTCCGGATGGGTAAACT	543
Db			
Qy	481	GGGTCCCATTTGGCCCTTTCTCTTTGGCGGGCACTGTGCTGGAAAGCGTAGACA	540
Db			
Qy	544	GGGTCCCATTTGTGCCCTTTTCTCTTTGGCGGGCACTGTGCTGGAAAGCGTAGACA	603
Db			
Qy	541	AGGAGATGCAAGTATTGGTGAATCGGATTCGAAGTTGGATGGCCACCTTACTGTAATGAC	600
Db			
Qy	604	AGGAGATGCAAGTATTGGTGAATCGGATTCGAAGTTGGATGGCCACCTTACTGTAATGAC	663
Db			
Qy	601	ACCTAGAGCTTGGATCAGAGAAACCGCGCTGGGACATTTTGTGGATCTCTACGGGA	660
Db			
Qy	664	ACCTAGAGCTTGGATCAGAGAAACCGCGCTGGGATCTTTTGTGGAACTCTATGGGA	723
Db			
Qy	661	ACAATGACAGCCGAGAGCCGGAAGAGCCAGGAGCGTTTCAACCGCTGGTCTCTGACGG	720
Db			
Qy	724	ACAATGACAGCCGAGAGCCGGAAGAGCCAGGAAACGCTTCAACCGCTGGTCTCTGACGG	783
Db			
Qy	721	GCATGACTGGCTGGTGTAGTTCTTGCTGGGCTCACTCTTCAGTCCGGAAGTACCAGACA	780
Db			

```
Db 184 TTTCCAGAAAGGATACAGCTGGAGTCACTTTAGTGTGGAAGAGAAAGGACTGAGG 243
Qy 181 CCCAGAGAAAGTGAACAGAAAGGAGAGACCCAGTGGCCATCAATGGCAACCCATCCT 240
Db 244 CCCAGAGAGGAGTGAATCGAGATGGAGACCCAGTGGCCATCAATGGCAACCCATCCT 303
Qy 241 GGCACCTGGCGATAGACCCCGGTGTAATGAGCCACTGGCCACAGCAGAGCTTTGGATG 300
Db 304 GGCACCTGGCAGACAGACCCCGGTGTAATGAGCCACTGGCCACAGCAGAGCTTTGGATG 363
Qy 301 CGCGGAGGTAAATCCCATGTCAGCAGTGAAGAGCGCTGAGAGAGCTGCGCATAGT 360
Db 423 CCCGGAGGTGATCCCATGTCAGCAGTGAAGAGCGCTGAGGAGGACAGCGCAGGT 423
Qy 361 TTGAATCGCGTACCGGAGAGCATTTCACTGATCTAAACATCCAGCTTCATATAACCCAG 420
Db 424 TTGAATCGCGTACCGGAGAGCATTTCACTGATCTAAACATCCAGCTTCATATAACCCAG 483
Qy 421 GGACAGCATATCAGAGCTTTGAACAGGTAGTGAATGAATCACTTTTCGGGATGGGTAACT 480
Db 484 GGACAGCATATCAGAGCTTTGAACAGGTAGTGAATGAATCACTTTTCGGGATGGGTAACT 543
Qy 481 GGGGTGCGATTTGGCCCTTTTCTCTTTCGGGGGCACTGTGGTGAAGAGCGTAGACA 540
Db 544 GGGGTGCGATTTGGCCCTTTTCTCTTTCGGGGGCACTGTGGTGAAGAGCGTAGACA 603
Qy 541 AGGAGATCGAGTATTTGTGAGTGGATTTGCAAGTTGGATGGCCACCTTACCTGAATGACC 600
Db 604 AGGAGATCGAGTATTTGTGAGTGGATTTGCAAGTTGGATGGCCACCTTACCTGAATGACC 663
Qy 601 ACCTAGAGCTTGGATTCAGAGAACCGCGGCTGGGACACTTTTGTGATCTCTACCGGA 660
Db 664 ACCTAGAGCTTGGATTCAGAGAACCGCGGCTGGGATACCTTTTGTGAACTCTATGGGA 723
Qy 661 ACAATGACAGCCGAGAGCGGAAAGCGCAGGAGCGTTTCAACCGCTGTTCTTGAGG 720
Db 724 ACAATGACAGCCGAGAGCGGAAAGCGCAGGAGCGTTTCAACCGCTGTTCTTGAGG 783
Qy 721 GCATGACTGTGGCTGTGTAGTTCCTGCTGGGCTCACTCTCAGTGGGAAGTAGACAGACA 780
Db 784 GCATGACTGTGGCTGTGTAGTTCCTGCTGGGCTCACTCTCAGTGGGAAGTAGACAGACA 843
Qy 781 CTGACCGTCCACTCACTCTCACTCCACTTGGCCCAACCAACTCTCTTTACGCC 840
Db 844 CTGACCATCCACTCACTCTCCACTCCCTCTCTGCTCCACCATCTCTCGTCCAGCC 903
Qy 841 ACCATTGCTACGAGAGAAC 861
Db 904 GCCATTGCCACGAGAGAAC 924
```

## RESULT 6

```
US-10-072-830-3
; Sequence 3, Application US/10072830
; Publication No. US20030103945A1
; GENERAL INFORMATION:
; APPLICANT: CHEN, DONG FENG
; APPLICANT: HUANG, XIZHONG
; APPLICANT: CHEN, GUANG
; APPLICANT: MANJI, HUSSEINI K.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR STIMULATING AXON
; REGENERATION AND PREVENTING NEURONAL CELL DEGENERATION
; FILE REFERENCE: ERW-105.01
; CURRENT APPLICATION NUMBER: US/10/072,830
; PRIORITY FILING DATE: 2002-02-08
; PRIORITY FILING DATE: 60/267,832
; PRIORITY FILING DATE: 2001-02-09
; PRIORITY FILING DATE: 60/272,617
; PRIORITY FILING DATE: 2001-03-01
; PRIORITY FILING DATE: 60/289,990
; PRIORITY FILING DATE: 2001-05-10
; NUMBER OF SEQ ID NOS: 8
```

```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 926
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (135)..(833)
US-10-072-830-3
```

```
Query Match 42.4%; Score 737.8; DB 15; Length 926;
Best Local Similarity 91.1%; Pred. No. 2e-221;
Matches 784; Conservative 0; Mismatches 77; Indels 0; Gaps 0;
```

```
Qy 1 CACAGAGACAGACCAGTGTGAGTGTGTTTGGACAATGGACTGTGTTGAGCCCATCT 60
Db 64 CAGAGACGAGACTCAGTGTGAGTGTGTTTGGACAATGGACTGTGTTGAGCCCATCT 123
Qy 61 CTATTATATAAATCTCTCAGAGCAACCGGGAGCTGGTGTGACTTTCTCTCTACAGC 120
Db 124 CTATTATATAAATCTCTCAGAGCAACCGGGAGCTGGTGTGACTTTCTCTCTACAGC 183
Qy 121 TCTCCAGAAAGGATACAGCTGGAGTCACTTTAGCGATGTGAAAGAGAAAGGACTGAAG 180
Db 184 TTTCCAGAAAGGATACAGCTGGAGTCACTTTAGCGATGTGAAAGAGAAAGGACTGAAG 243
Qy 181 CCCAGAGAAATGAAACCAAGAGGAGAGACCCCGAGTGCCATCAATGGCAACCCATCCT 240
Db 244 CCCAGAGAGGACTGAATCGAGATGGAGACCCCGAGTGCCATCAATGGCAACCCATCCT 303
Qy 241 GGCACCTGGCGATAGACCCCGGTGAAATGGAGCCACTGGCCACAGCAGCAGTTGGATG 300
Db 304 GGCACCTGGCAGACAGACCCCGGTGAAATGGAGCCACTGGCCACAGCAGCAGTTGGATG 363
Qy 301 CGCGGAGGTAAATCCCATGGCAGCAGTGAAGAGCGCTGAGAGAGCTGGCGATGAGT 360
Db 364 CCCGGAGGTGATCCCATGGCAGCAGTGAAGAGCGCTGAGGAGGAGCGGAGCAGT 423
Qy 361 TTGAACCTCGGTACCGGAGAGCATTTCACTGATCTTAACATCCAGCTTCATATAACCCAG 420
Db 424 TTGAACCTCGGTACCGGAGAGCATTTCACTGATCTTAACATCCAGCTTCATATAACCCAG 483
Qy 421 GGACAGCATATCAGAGCTTTGAACAGGTAGTGAATGAATCTTTTCGGGATGGGTAACT 480
Db 484 GGACAGCATATCAGAGCTTTGAACAGGTAGTGAATGAATCTTTTCGGGATGGGTAACT 543
Qy 481 GGGGTGCGATTTGGCCCTTTTCTCTTTCGGGGGCACTGTGGTGAAGAGCGTAGACA 540
Db 544 GGGGTGCGATTTGGCCCTTTTCTCTTTCGGGGGCACTGTGGTGAAGAGCGTAGACA 603
Qy 541 AGGAGATCGAGTATTTGTGAGTGGATTTGCAAGTTGGATGGCCACCTTACCTGAATGACC 600
Db 604 AGGAGATCGAGTATTTGTGAGTGGATTTGCAAGTTGGATGGCCACCTTACCTGAATGACC 663
Qy 601 ACCTAGAGCTTGGATTCAGAGAACCGCGGCTGGGACACTTTTGTGATCTCTACGGGA 660
Db 664 ACCTAGAGCTTGGATTCAGAGAACCGCGGCTGGGATACCTTTTGTGAACTCTATGGGA 723
Qy 661 ACAATGACAGCCGAGAGCGGAAAGCGCAGGAGCGTTTCAACCGCTGTTCTTGAGCG 720
Db 724 ACAATGACAGCCGAGAGCGGAAAGCGCAGGAGCGTTTCAACCGCTGTTCTTGAGCG 783
Qy 721 GCATGACTGTGGCTGTGTAGTTCCTGCTGGGCTCACTCTCAGTGGGAAGTAGACAGACA 780
Db 784 GCATGACTGTGGCTGTGTAGTTCCTGCTGGGCTCACTCTCAGTGGGAAGTAGACAGACA 843
Qy 781 CTGACCGTCCACTCACTCTCACTCCACTTGGCCCAACCAACTCTCTTTACGCC 840
Db 844 CTGACCATCCACTCACTCTCCACTCCCTCTCTGCTCCACCATCTCTCGTCCAGCC 903
Qy 841 ACCATTGCTACGAGAGAAC 861
Db 904 GCCATTGCCACGAGAGAAC 924
```

## RESULT 7

US-10-302-262-1

; Sequence 1, Application US/10302262  
; Publication No. US20030191300A1  
; GENERAL INFORMATION:  
; APPLICANT: Bennett, C. Frank  
; APPLICANT: Dean, Nicholas M.  
; APPLICANT: Monia, Brett P.  
; APPLICANT: Nickoloff, Brian J.  
; APPLICANT: Zhang, QingQing  
; TITLE OF INVENTION: Antisense Modulation of bcl-x Expression  
; FILE REFERENCE: ISPH-0528  
; CURRENT APPLICATION NUMBER: US/10/302,262  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/734,846  
; PRIOR FILING DATE: 2000-12-12  
; PRIOR APPLICATION NUMBER: 09/277,020  
; PRIOR FILING DATE: 1998-03-26  
; PRIOR APPLICATION NUMBER: 09/167,921  
; PRIOR FILING DATE: 1998-10-07  
; PRIOR APPLICATION NUMBER: 09/323,743  
; PRIOR FILING DATE: 1999-06-02  
; NUMBER OF SEQ ID NOS: 74  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 926  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (135)..(836)  
; DATABASE ACCESSION NUMBER: L20121 Genbank  
; DATABASE ENTRY DATE: 1994-07-26  
US-10-302-262-1

Query Match 42.4%; Score 737.8; DB 16; Length 926;

Best Local Similarity 91.1%; Pred. No. 2e-221; Indels 0; Gaps 0;

Matches 784; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

Qy 1 CACAGCAGACCCAGTCAGTCAGCAGAGTGTGGACAAATGGAGCTGGTTCAGGCCATCT 60  
Db |||||  
Qy 64 CAGAGCAGACTCAGTCAGTCAGCAGAGTGTGGACAAATGGAGCTGGTTCAGGCCATCT 123  
Db |||||  
Qy 61 CTATTATATAAATGTCCTCAGAGCAACCGGAGCTGGTGGTTCCTCTCTACAAAGC 120  
Db |||||  
Qy 124 CTATTATATAAATGTCCTCAGAGCAACCGGAGCTGGTGGTTCCTCTCTACAAAGC 183  
Qy 121 TCTCCAGAAAGGATACAGCTGGAGTGTAGCGATGTGGAAGAAACAGGACTGAAG 180  
Db |||||  
Qy 184 TTTCCAGAAAGGATACAGCTGGAGTGTAGCGATGTGGAAGAAACAGGACTGAAG 243  
Db |||||  
Qy 181 CCCAGAAAGAACTGAACAGAAAGGAGAGCCCGAGTGCCATCAATGGCAACCCATCCT 240  
Db |||||  
Qy 244 CCCCAGAAAGGACTGAATCGGAGATGGAGAGCCCGAGTGCCATCAATGGCAACCCATCCT 303  
Qy 241 GGCACTGGCGGATAGCCCGCGTGAATGAGGCCACTGGCCACAGCAGCAGTTGGATG 300  
Db |||||  
Qy 304 GGCACTGGCAGACAGCCCGCGTGAATGAGGCCACTGGCCACAGCAGCAGTTGGATG 363  
Qy 301 CGCGGGAGTATCCCATGGCAGCAGTGAAGCAGCGCTGAGAGGCGCTGGCGATGAGT 360  
Db |||||  
Qy 364 CCGGGAGGTGATCCCATGGCAGCAGTGAAGCAGCGCTGAGAGGAGCGCGGACGAGT 423  
Qy 361 TTGAACCTGGGTACCGGAGAGCATTCAGTGATCTAAACATCCCGCTTCATATAACCCAG 420  
Db |||||  
Qy 424 TTGAACCTGGGTACCGGAGAGCATTCAGTGATCTAAACATCCCGCTTCATATAACCCAG 483  
Qy 421 GGACAGCATATCAGAGCTTTGAAACAGGTAGTGAATGAATCTTTTCGGGATGGGGTAACT 480  
Db |||||  
Qy 484 GGACAGCATATCAGAGCTTTGAAACAGGTAGTGAATGAATCTTTTCGGGATGGGGTAACT 543  
Db |||||

Qy 481 GGGTCGCATTGTGGCCCTTCTCTCTTTGGCGGGGCACTGTGCGTGGAAAGCGTAGACA 540  
Db |||||  
Qy 544 GGGTCGCATTGTGGCCCTTCTCTCTTTGGCGGGGCACTGTGCGTGGAAAGCGTAGACA 603  
Db |||||  
Qy 541 AGGAGATGCAGGTATTGGTGGATCGGATTCGAAAGTGGATGGCCACCTTACCTGAATGACC 600  
Db |||||  
Qy 604 AGGAGATGCAGGTATTGGTGGATCGGATTCGAAAGTGGATGGCCACCTTACCTGAATGACC 663  
Db |||||  
Qy 601 ACCTAGAGCCTTGGATCCAGGAGAACGGCGGCTGGGACACTTTTGTGGATCTCTACGGGA 660  
Db |||||  
Qy 664 ACCTAGAGCCTTGGATCCAGGAGAACGGCGGCTGGGATACCTTTTGTGGAACTCTATGGGA 723  
Db |||||  
Qy 661 ACATGACAGCAGCGAGAGCCGAAAGCCAGGAGCGTTTCAACGCGTGGTTCCTGACGG 720  
Db |||||  
Qy 724 ACATGACAGCAGCGAGAGCCGAAAGCCAGGAGCGTTTCAACGCGTGGTTCCTGACGG 783  
Db |||||  
Qy 721 GCATGACTGTGGCTGGTGTAGTTCTGCTGGGCTCACTCTTCAGTCGGAAGTGACCAACA 780  
Db |||||  
Qy 784 GCATGACTGTGGCGCGGTGTTCTGCTGGGCTCACTCTTCAGTCGGAAGTGACCAACA 843  
Qy 781 CTGACCGTCCACTCACTCTCACTCCACCTCCACCTGCGCCCGCCACCACTCTCTCTCAGCC 840  
Db |||||  
Qy 844 CTGACCATCCACTCTACCTCTCCACCTCCACCTGCGCCCGCCACCACTCTCTCTCAGCC 903  
Qy 841 ACCATTGCTACCAAGGAGAAC 861  
Db |||||  
Qy 904 GCCATTGCCACCAAGGAGAAC 924  
Db |||||

## RESULT 8

US-10-402-017-1

; Sequence 1, Application US/10402017  
; Publication No. US20030219871A1  
; GENERAL INFORMATION:  
; APPLICANT: Barbara ENENKEL, Heiko MEENTS and Martin FUSSENEGGER  
; TITLE OF INVENTION: Host cells having improved survival properties and methods to gen  
; TITLE OF INVENTION: such cells  
; FILE REFERENCE: Case 1/1314  
; CURRENT APPLICATION NUMBER: US/10/402,017  
; CURRENT FILING DATE: 2003-03-28  
; PRIOR APPLICATION NUMBER: US 60/369,307  
; PRIOR APPLICATION NUMBER: April 2, 2002  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 926  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-402-017-1

Query Match 42.4%; Score 737.8; DB 17; Length 926;

Best Local Similarity 91.1%; Pred. No. 2e-221; Indels 0; Gaps 0;

Matches 784; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

Qy 1 CACAGCAGACCCAGTCAGTCAGCAGAGTGTGGACAAATGGAGCTGGTTCAGGCCATCT 60  
Db |||||  
Qy 64 CAGAGCAGACTCAGTCAGTCAGCAGAGTGTGGACAAATGGAGCTGGTTCAGGCCATCT 123  
Db |||||  
Qy 61 CTATTATATAAATGTCCTCAGAGCAACCGGAGCTGGTGGTTCCTCTCTACAAAGC 120  
Db |||||  
Qy 124 CTATTATATAAATGTCCTCAGAGCAACCGGAGCTGGTGGTTCCTCTCTACAAAGC 183  
Qy 121 TCTCCAGAAAGGATACAGCTGGAGTGTAGCGATGTGGAAGAAACAGGACTGAAG 180  
Db |||||  
Qy 184 TTTCCAGAAAGGATACAGCTGGAGTGTAGCGATGTGGAAGAAACAGGACTGAAG 243  
Db |||||  
Qy 181 CCCAGAAAGAACTGAACAGAAAGGAGAGCCCGAGTGCCATCAATGGCAACCCATCCT 240  
Db |||||  
Qy 244 CCCCAGAAAGGACTGAATCGGAGATGGAGAGCCCGAGTGCCATCAATGGCAACCCATCCT 303  
Qy 241 GGCACTGGCGGATAGCCCGCGTGAATGAGGCCACTGGCCACAGCAGCAGTTGGATG 300  
Db |||||  
Qy 304 GGCACTGGCAGACAGCCCGCGTGAATGAGGCCACTGGCCACAGCAGCAGTTGGATG 363  
Db |||||



```

301 CGCGGAGGTAATCCCATGCGACAGTGAAGCAAGCGCTGAGAGAGCGCTGGCGATGAGT 360
364 CCCGGGAGGTGATCCCATGCGACAGTGAAGCAAGCGCTGAGAGAGCGCTGGCGATGAGT 423
361 TTGAACGCGGTACCGGAGAGCAATTCAGTGATCTAACTCCAGCTTCATATACCCGAG 420
424 TTGAACGCGGTACCGGAGAGCAATTCAGTGATCTAACTCCAGCTTCATATACCCGAG 483
421 GGACAGCATATCAGAGCTTTGAACAGGTAGTGAATGAATCTTTTCGGGATGGGTAAGT 480
484 GGACAGCATATCAGAGCTTTGAACAGGTAGTGAATGAATCTTTTCGGGATGGGTAAGT 543
481 GGGTGCAGATTGTGGCTTCTTCTTTGCGGGGCACTGTGCGTGGAAAGCGTAGACA 540
544 GGGTGCAGATTGTGGCTTCTTCTTTGCGGGGCACTGTGCGTGGAAAGCGTAGACA 603
541 AGGAGATCGAGTATGCGTGGAGTGGATTGCAAGTTGGATGCGCACCTACCTGAATGACC 600
604 AGGAGATCGAGTATGCGTGGAGTGGATTGCAAGTTGGATGCGCACCTACCTGAATGACC 663
601 ACCTAGAGCCTTGGATCCAGAGAGAACCGCGGCTGGGACACTTTTGTGGATCTCTACGGGA 660
664 ACCTAGAGCCTTGGATCCAGAGAGAACCGCGGCTGGGATCTTTTGTGGATCTCTATGGGA 723
661 ACAATGACAGCGCGAGAGCGGAAAGCGGAGAGCGGTTTCAACCGCTGGTTCCTGACGG 720
724 ACAATGACAGCGCGAGAGCGGAAAGCGGAGAGCGGTTTCAACCGCTGGTTCCTGACGG 783
721 GCATGACTGTGGCTGTGATGTTCTGCTGGGCTCACTTTCAGTGGAGTGCAGACACA 780
784 GCATGACTGTGGCTGTGATGTTCTGCTGGGCTCACTTTCAGTGGAGTGCAGACACA 843
781 CTGACCGTCCACTCACTCTCACTCCACCTTCCACCTTCCACCTTCTGCTCCACCATCTCTCCAGCC 840
844 CTGACCGTCCACTCACTCTCACTCCACCTTCCACCTTCCACCTTCTGCTCCACCATCTCTCCAGCC 903
841 ACCATTGCTACCGAGAGAAC 861
904 GCCATTGCCACGAGAGAAC 924

```

RESULT 9

```

US-10-641-643-1430
; Sequence 1430, Application US/10641643
; Publication No. US20040077003A1
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
;              Susan G. Stuart
;              Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL
; GENE EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESS: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/641,643
; FILING DATE: 14-Aug-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: <Unknown>
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:

```

```

; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1430:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 926 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: 9510900
; SEQUENCE DESCRIPTION: SEQ ID NO: 1430 :
US-10-641-643-1430

Query Match      42.4%; Score 737.8; DB 17; Length 926;
Best Local Similarity 91.1%; Pred. No. 2e-221;
Matches 784; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 1 CACAGACGACACCCAGTGTAGTGACGAGTGTGTTGGACAATGGACTGGTTCAGGCCATCT 60
DB 64 CAGAGACGAGACTCAGTGAGTGACGAGTGTGTTGGACAATGGACTGGTTCAGGCCATCT 123
QY 61 CTATTATAAANAATGCTCAGAGCAACCGGAGCTGGTGGTGGACTTTCTCTCTACAGC 120
DB 124 CTATTATAAANAATGCTCAGAGCAACCGGAGCTGGTGGTGGACTTTCTCTCTACAGC 183
QY 121 TCTCCCAAGAAAGATACAGCTGGAGTCACTTTAGCGATGTGAAAGAGAACAGGACTGAAG 180
DB 184 TTTCCCAAGAAAGATACAGCTGGAGTCACTTTAGCGATGTGAAAGAGAACAGGACTGAAG 243
QY 181 CCCCAGAGAAACTGAACACGAGAGGAGAGACCCCGAGTGCCATCAATGGCAACCCATCCT 240
DB 244 CCCCAGAGAGGACTGAATCGAGATGGAGACCCCGAGTGCCATCAATGGCAACCCATCCT 303
QY 241 GGCACCTGGCGGATAGCCCGCGGTGAATGAGAGCACCTGGCCACAGCAGCAGTGTGGATG 300
DB 304 GGCACCTGGCAGACAGACCCCGCGGTGAATGAGAGCACCTGGCCACAGCAGCAGTGTGGATG 363
QY 301 CGCGGAGGTAATCCCATGCGACAGTGAAGCAAGCGCTGAGAGAGCGCTGGCGATGAGT 360
DB 364 CCCGGGAGGTGATCCCATGCGACAGTGAAGCAAGCGCTGAGAGAGCGCAGCGAGT 423
QY 361 TTGAACGCGGTACCGGAGAGCAATTCAGTGATCTAACTCCAGCTTCATATACCCGAG 420
DB 424 TTGAACGCGGTACCGGAGAGCAATTCAGTGATCTAACTCCAGCTTCATATACCCGAG 483
QY 421 GGACAGCATATCAGAGCTTTGAACAGGTAGTGAATGAATCTTTTCGGGATGGGTAAGT 480
DB 484 GGACAGCATATCAGAGCTTTGAACAGGTAGTGAATGAATCTTTTCGGGATGGGTAAGT 543
QY 481 GGGTGCAGATTGTGGCTTCTTCTTTGCGGGGCACTGTGCGTGGAAAGCGTAGACA 540
DB 544 GGGTGCAGATTGTGGCTTCTTCTTTGCGGGGCACTGTGCGTGGAAAGCGTAGACA 603
QY 541 AGGAGATCGAGTATGCGTGGAGTGGATTGCAAGTTGGATGCGCACCTACCTGAATGACC 600
DB 604 AGGAGATCGAGTATGCGTGGAGTGGATTGCAAGTTGGATGCGCACCTACCTGAATGACC 663
QY 601 ACCTAGAGCCTTGGATCCAGAGAGAACCGCGGCTGGGACACTTTTGTGGATCTCTACGGGA 660
DB 664 ACCTAGAGCCTTGGATCCAGAGAGAACCGCGGCTGGGATCTTTTGTGGATCTCTATGGGA 723
QY 661 ACAATGACAGCGCGAGAGCGGAAAGCGGAGAGCGGTTTCAACCGCTGGTTCCTGACGG 720
DB 724 ACAATGACAGCGCGAGAGCGGAAAGCGGAGAGCGGTTTCAACCGCTGGTTCCTGACGG 783
QY 721 GCATGACTGTGGCTGTGATGTTCTGCTGGGCTCACTTTCAGTGGAGTGCAGACACA 780
DB 784 GCATGACTGTGGCTGTGATGTTCTGCTGGGCTCACTTTCAGTGGAGTGCAGACACA 843

```







```
QY 541 AGGAGATCCAGGTATTGGTGAGTCGGATTGCAAGTTGGATGGCCACCTACCTGTAATGACC 600
Db -----740
QY 601 ACCTAGAGCCTTGGATCCAGGAGAACGGCGCTGGGACACTTTTGTGATCTCTACGGGA 660
Db -----766
QY 661 ACAATGACAGCCGAGAGCCGGAAGAGCGTTTCAACCGCTGGTTCTCTGACGG 720
Db -----826
QY 721 GCATGACTGTGGCTGGTGTAGTTCTGTGGGCTCACTTTCAGTCGGAAGTACCAAGACA 780
Db -----886
QY 781 CTGACCGTCCACTCACCTCTACCTCCACCTTGCCGCCACACCAACTCTCTCTTCAGCC 840
Db -----946
QY 841 ACCATTGCTACAGGAGAACACTACATGCAACTCACGCCCTTCCCCTATTATAGGGTT 900
Db -----1006
QY 901 GGGCTTAGACGGAGTCCCTCGAGTTAGCTTTCTAGAACTACACAGCTTCTGTGAAGC 960
Db -----1064
QY 961 CACCTTCCCCCACAATCTCAGTTCCTTGGGCTCAAACTCACAAGTTTTCCTTCAGAT 1020
Db -----1116
QY 1021 CAGCTCTTGGAGGTGCGAGAGTGGGAAGGGGTGTCTAGAGGG--AGAAGAGCCTGC 1078
Db -----1176
QY 1079 CTTGTTGGTGGACCTGATTAACCTCAGCCCTGAGCCCTCTCGGGAATGCTTTTCTGGCAGGAGC 1138
Db -----1234
QY 1139 TGGGAGCTCTTAACCTCTTCCCC--CAGAGACTAGATTGCTTGTGTTTGTGATGTG 1194
Db -----1294
QY 1195 TGTGGCTCAGAAATGATTCATTTTCCCCCTCTGCTGTCTGCTGGGGGGCTCTCTCTTC 1254
Db -----1354
QY 1255 CCATCTCCACCCCTTGGATTAACAGAGCCATTGAGTGAGGTGTTTATGACCTTTTGACTA 1314
Db -----1398
QY 1315 ACTAAATATGAGCTGTGGGATAACAGGCAAGGACC-----TCTCTCCACCTGT 1368
Db -----1458
QY 1369 GGCCTGGGCA-AGCCCCCACTCTGTGCTGAATGTTCTCTGAGGCTCTGGCTAGAGTC 1427
Db -----1518
QY 1428 CAGCCCCCAGGAGTGGGATAAGTTCGGAAGTCCACCTTCGCGAGAGCCTGA 1484
Db -----1578
QY 1485 CGGGCTCTTGGGCTTAGACACACCCAGATCTTCTCCACCCCTCCCTGGCTCCATGG 1544
Db -----1629
QY 1545 TGACATGACTGAGGACCACTGGGCCACGCTAGGTGGCCAGAGCTGTTAATGACTT 1604
Db -----1689
```

```
QY 1605 CAGCTGCTCACTTCTCTGCAAGATCAGCCTGTGCACTTTTGGTCTGGTCTGGCCACA 1664
Db -----1746
QY 1665 GGGTCCAGGAGTCTGGCCTTAGCCAGAACTGAGAGNAAGCTTACA-GCGCAGCTATGG 1723
Db -----1806
QY 1747 CATGGTGGTTTCAGGGGACTCAGCCCTGAGGTGAAGGGAGCTATCAGGAACAGCTATGG 1806
Db -----1825
QY 1724 GAGCCCTGGGGCTTCCCT 1742
Db -----1825
RESULT 15
US-10-388-934-539
; Sequence 539, Application US/10388934
; Publication No. US20040005547A1
; GENERAL INFORMATION:
; APPLICANT: Boess, Franziska
; APPLICANT: Suter-Dick, Laura
; APPLICANT: Wolf, Detlef
; TITLE OF INVENTION: BIOMARKERS AND EXPRESSION PROFILES FOR TOXICOLOGY
; FILE REFERENCE: 21199
; CURRENT APPLICATION NUMBER: US/10/388,934
; CURRENT FILING DATE: 2003-03-14
; PRIOR APPLICATION NUMBER: 02005336.9
; PRIOR FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 02015657.6
; PRIOR FILING DATE: 2002-07-17
; NUMBER OF SEQ ID NOS: 862
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 539
; LENGTH: 1748
; TYPE: DNA
; ORGANISM: Rattus norvegicus (No. US20040005547Alway rat)
US-10-388-934-539
```

```
Query Match 36.5%; Score 636; DB 17; Length 1748;
Best Local Similarity 100.0%; Pred. No. 3.1e-189;
Matches 636; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CACAGACGACACCCAGTGTGACAGGTGTTTGGCAATGGACTGGTTCAGGCCATCT 60
Db -----152
QY 61 CTATTATAAAATGTCTCAGAGCAACCGGAGCTGGTGTGACTTTCTCTCTACAAGC 120
Db -----212
QY 121 TCTCCCAAGAAAGATACAGCTGGAGTCAGTTTACGATGTGCAAGAGAACAGGACTGAAG 180
Db -----272
QY 181 CCCCAGAGAACTACACCAAGAGGAGACCCCGCTGATGAGCCACTGGCCACAGCAGCTTGAATG 240
Db -----332
QY 241 GGCACCTGGCGATAGCCCGCTGATGAGCCACTGGCCACAGCAGCAGCTTGAATG 300
Db -----392
QY 301 CGCGGAGGTAATCCCATGCGCAGCTGAAGCAAGCGCTGAGAGAGGCTGGCGATGAGT 360
Db -----452
QY 361 TTGAACCTGGGTACCGGAGAGCATTCAGTGTATTAACATCCAGCTTCATATAACCCAG 420
Db -----512
QY 421 GGACAGCATATCAGAGCTTTGAACAGGTAGTGAATGAATCTTTTCGGGATGGGGTAACT 480
Db -----572
```

Qy	481	GGGTCGCATTGTGGCCCTTCTTCTCCTTTGGGGGGCACTGTGGCTGGAAAGCGTAGACA	540
Db	573	GGGTCGCATTGTGGCCCTTCTTCTCCTTTGGGGGGCACTGTGGCTGGAAAGCGTAGACA	632
Qy	541	AGGAGATGCAGGTATTGTGAGTCGGATTGCAAGTTGGATGGCCACCTACCTGAATGACC	600
Db	633	AGGAGATGCAGGTATTGTGAGTCGGATTGCAAGTTGGATGGCCACCTACCTGAATGACC	692
Qy	601	ACCTAGAGCCCTTGGATCCAGGAGAACGGCGGCTGGG	636
Db	693	ACCTAGAGCCCTTGGATCCAGGAGAACGGCGGCTGGG	728

Search completed: March 31, 2005, 06:13:03  
Job time : 6364 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 31, 2005, 00:31:42 ; Search time 5241 Seconds  
(without alignments)  
12651.776 Million cell updates/sec

Title: US-10-049-822A-1  
Perfect score: 1742  
Sequence: 1 cacagagcagaccagtgag.....ggagccctgggggttcct 1742

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST.\*  
1: gb\_est1.\*  
2: gb\_est2.\*  
3: gb\_hic.\*  
4: gb\_est3.\*  
5: gb\_est4.\*  
6: gb\_est5.\*  
7: gb\_est6.\*  
8: gb\_gss1.\*  
9: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	763	43.8	894	5	BQ935140 AGENCOURT
2	742.6	42.6	767	7	CK483979 AGENCOURT
3	727.4	41.8	1087	5	BUI56982 AGENCOURT
4	727	41.7	919	5	BQ948029 AGENCOURT
5	726.4	41.7	878	5	BQ888053 AGENCOURT
6	722.4	41.5	997	5	BQ052167 AGENCOURT
7	716.2	41.1	989	5	BX441570 AGENCOURT
8	714.8	41.0	1002	4	BM549716 AGENCOURT
9	707.2	40.6	758	7	CO396360 AGENCOURT
10	707	40.6	1134	5	BX359456 AGENCOURT
11	704.4	40.4	798	7	CK000319 AGENCOURT
12	696.2	40.0	922	5	BQ924197 AGENCOURT
13	693.6	39.8	732	5	BQ053776 AGENCOURT
14	693.4	39.8	1035	4	BI409987 AGENCOURT
15	691.8	39.7	826	7	CK482364 AGENCOURT
16	690	39.6	1045	4	BM811850 AGENCOURT
17	686.2	39.4	724	7	CF730725 AGENCOURT
18	685.2	39.3	735	5	BU053967 AGENCOURT
19	682.2	39.2	720	5	BU709510 AGENCOURT
20	680.8	39.1	798	7	CK481691 AGENCOURT
21	680.2	39.0	719	5	BQ443141 AGENCOURT
22	676.4	38.8	721	5	BU703310 AGENCOURT
23	668.4	38.4	913	4	BG244406 AGENCOURT
24	663.2	38.1	700	6	CA751645 AGENCOURT

25	663.2	38.1	752	5	BU614401	UI-M-EVO-
26	657.2	37.7	872	5	BUI68531	AGENCOURT
c	655.2	37.6	862	4	BG973530	602842233
28	648.2	37.2	685	7	CF173595	B0926B04-
29	646.6	37.1	685	7	CF172085	B0902C04-
30	646.2	37.1	683	5	BQ178703	UI-M-EVO-
31	645.2	37.0	963	5	BQ643540	AGENCOURT
32	642.8	36.9	1028	5	BM928004	AGENCOURT
33	641	36.8	1010	4	BM556041	AGENCOURT
34	636.2	36.5	676	5	BU709215	UI-M-EVO-
35	635.4	36.5	809	7	CK128475	AGENCOURT
36	627.8	36.0	972	5	BQ646866	AGENCOURT
37	626.4	36.0	923	5	BQ927754	AGENCOURT
38	624.6	35.9	773	7	CV111170	AGENCOURT
39	623.8	35.8	843	5	BQ962018	AGENCOURT
40	622.8	35.8	751	7	CK022949	AGENCOURT
41	618.6	35.5	819	7	CK601476	AGENCOURT
42	617.8	35.5	653	7	CF915537	B0981E03-
43	615.4	35.3	720	6	CD636470	56049223J
44	614.4	35.3	956	5	BQ648225	AGENCOURT
45	610	35.0	830	7	CF619432	AGENCOURT

## ALIGNMENTS

RESULT 1  
LOCUS BQ935140 894 bp mRNA linear EST 21-AUG-2002  
DEFINITION AGENCOURT 8980602 NCI\_CGAP\_Co24 Mus musculus cDNA clone IMAGE:6397177 5', mRNA sequence.  
ACCESSION BQ935140  
VERSION BQ935140.1 GI:22350523  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 894)  
TITLE NIH-MGC http://mgi.nci.nih.gov/  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabs-r@mail.nih.gov  
Tissue Procurement: The Cepko Laboratory  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov  
Plate: LLAM13894 row: k column: 02  
High quality sequence stop: 657.  
Location/Qualifiers  
1. 894  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:6397177"  
/lab\_host="DH10B (T1 phage-resistant)"  
/clone\_lib="NCI\_CGAP Co24"  
/note="Organ: colon; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.6 kb. Constructed by Life Technologies. Note: this is a NCI\_CGAP Library."

## FEATURES

source  
Query Match 43.8%; Score 763; DB 5; Length 894;  
Best Local Similarity 95.9%; Pred. No. 1e-197;  
Matches 804; Conservative 0; Mismatches 31; Indels 3; Gaps 2;  
QY 85 ACCGGGAGCTGGTGGTTGACTTCTCTCTACAGCTCTCCAGAAAGGATACAGCTGGA 144

Db 1 ACCGGGAGCTGGTGGTGCAGCTTCTCTCTCAAGCTTTCCAGAAAGATACAGCTGGA 60  
Qy 145 GTCAGTTAGCGATGTCGAGAGACAGGACTGAAGCCCGCAGAGAACTGAACAGAAA 204  
Db 61 GTCAGTTAGTGTGATGCGAAGAGATAGGACTGAGGCCCGCAGAGAACTGAAGCAGAGA 120  
Qy 205 GGGAGACCCCGAGTCCCATCAATGCAACCCATCTCTGACCTGCGACCTGCGGATAGCCCGGG 264  
Db 121 GGGAGACCCCGAGTCCCATCAATGCAACCCATCTCTGACCTGCGGATAGCCCGGG 180  
Qy 265 TGAATGGAGCCACTGGCCACAGCAGCTTTGGATGCGCGGAGGATTAATCCCATGGCAG 324  
Db 181 TGAATGGAGCCACTGGCCACAGCAGCTTTGGATGCGCGGAGGATTAATCCCATGGCAG 240  
Qy 325 CAGTGAACCAAGCGCTGAGAGAGGCTGGGATGAGTTTGAATGCGGTTACCGGAGAGAT 384  
Db 241 CAGTGAACCAAGCGCTGAGAGAGGCGGATGAGTTTGAATGCGGTTACCGGAGAGCGT 300  
Qy 385 TCAGTGATCTTAACATCCAGCTTTCATATAACCCAGGACAGCATATCAGAGCTTTGAAC 444  
Db 301 TCAGTGATCTTAACATCCAGCTTTCATATAACCCAGGACAGCGGTATCAGAGCTTTGAGC 360  
Qy 445 AGGTAGTGAATGAACCTTTTCGGGATGGGTAAACTGGGGTGGCATTTGGCCCTTTCT 504  
Db 361 AGGTAGTGAATGAACCTTTTCGGGATGGGTAAACTGGGGTGGCATTTGGCCCTTTCT 420  
Qy 505 CTTTGGCGGGGACCTGTGCTGGAAGCGGTAGACAAGAGATGAGATGATTTGGTGAATC 564  
Db 421 CTTTGGCGGGGACCTGTGCTGGAAGCGGTAGACAAGAGATGAGATGATTTGGTGAATC 480  
Qy 565 GGATTTGCAAGTTGGATGCGCCACTACCTGATGATGACCACTGAGCTTGGATCCAGGAGA 624  
Db 481 GGATTTGCAAGTTGGATGCGCCACTACCTGATGATGACCACTGAGCTTGGATCCAGGAGA 540  
Qy 625 ACGCGGCTGGGACACTTTTGTGGATCTCTACGGGAACAATGCAGCAGCCGAGAGCCGGA 684  
Db 541 ACGCGGCTGGGACACTTTTGTGGATCTCTACGGGAACAATGCAGCAGCCGAGAGCCGGA 600  
Qy 685 AAGGCAGAGGCTTTCAACCGCTGGTTCCTGACCGGGATGATGCTGCTGGTGTAGTTC 744  
Db 601 AAGGCAGAGGCTTTCAACCGCTGGTTCCTGACCGGGATGATGCTGCTGGTGTAGTTC 660  
Qy 745 TGCTGGGCTCACTCTTTCAGTGGAGTGCAGACACTGACCTGACCTCACTCACTCTCAC 804  
Db 661 TGCTGGGCTCACTCTTTCAGTGGAGTGCAGACACTGACCTGACCTCACTCACTCTCAC 720  
Qy 805 TCCCACTTGGCCCCCACCACCACTCTCTTTCAGCCACCATTTGCTTACCGAGAGAACCACT 864  
Db 721 TCCCACTTGGCCCCCACCACCACTCTCTTTCAGCCACCATTTGCTTACCGAGAGAACCACT 780  
Qy 865 ACATGCAACTACG--CCGCTTCCCTATATAGGGTTGGG--CTTAGAGGAGTCCCC 919  
Db 781 ACATGCAACTACGCCCCCTTGGCCCTATACAGGGTTGGGGCCCTTAGAGGAGTCCCC 838

RESULT 2  
LOCUS CK483979 767 bp mRNA linear EST 14-JAN-2004  
DEFINITION AGENCOURT 17635534 NIH MGC 235 Rattus norvegicus cDNA clone  
IMAGE:7110877 5', mRNA sequence.  
ACCESSION CK483979  
KEYWORDS CK483979.1 GI:40828327  
SOURCE EST.  
ORGANISM Rattus norvegicus (Norway rat)  
Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
REFERENCE 1 (bases 1 to 767)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)

COMMENT

Contact: Daniela S. Gerhard, Ph.D.  
Office of Cancer Genomics  
National Cancer Institute / NIH  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: cgabs-remail.nih.gov  
Tissue Procurement: Howard Jacobs  
cDNA Library Preparation: Express Genomics  
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM14978 row: 1 column: 11  
High quality sequence stop: 675.  
Location/Qualifiers  
1. .767  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10116"  
/clone="IMAGE:7110877"  
/tissue\_type="kidney, pooled"  
/lab\_host="DH10B Tona"  
/clone\_lib="NIH\_MGC 235"  
Site 2: NotI; RNA obtained from pooled kidney tissue from  
a mix of male and female animals at 8 wk old. Tissues were  
snap-frozen before RNA extraction and purification  
(TRI-reagent method). cDNA was primed using oligo-dT  
primer: 5'-pGACTAGTCTAGTCGCGGCGGCC(T)25-3' and  
cloned into the EcoRV/NotI sites of pExpress-1.  
Size-selection >1.4kb resulted in an average insert size  
of 2.2 kb. This primary library is non-normalized  
(normalized primary library is NIH\_MGC\_236) and was  
constructed by Express Genomics (Frederick, MD). Note:  
this is a NIH\_MGC library."

FEATURES  
source

Query Match 42.6%; Score 742.6; DB 7; Length 767;  
Best Local Similarity 99.2%; Pred. No. 3.9e-192;  
Matches 757; Conservative 0; Mismatches 2; Gaps 1;  
Qy 234 CCATCTCTGGCACCTGGCGGATAGCCCGCGGTGAATGGAGCCACTGGCCACAGCAGCAGT 293  
Db 1 CCATCTCTGGCACCTGGCGGATAGCCCGCGGTGAATGGAGCCACTGGCCACAGCAGCAGT 60  
Qy 294 TTGATGCGGGGAGGTAAATCCCATGCGCAGCATGTAAGCAAGCGCTGAGAGAGCTGGC 353  
Db 61 TTGATGCGGGGAGGTAAATCCCATGCGCAGCATGTAAGCAAGCGCTGAGAGAGCTGGC 120  
Qy 354 GATCAGTTTGAACCTGCGGTACCGGAGAGCATTCAGTCGATCTTAACATCCAGCTTCATATA 413  
Db 121 GATGAGTTTGAACCTGCGGTACCGGAGAGCATTCAGTCGATCTTAACATCCAGCTTCATATA 180  
Qy 414 ACCCCAGGACAGCATATCAGAGCTTTTGAACAGAGTAGTGAATGAATCTTTTCGGGATGGG 473  
Db 181 ACCCCAGGACAGCATATCAGAGCTTTTGAACAGAGTAGTGAATGAATCTTTTCGGGATGGG 240  
Qy 474 GTAAACTGGGGTGCATTTGGCCCTTCTCTTTTGGCGGGGCACTGTGCGTGGAAGC 533  
Db 241 GTAAACTGGGGTGCATTTGGCCCTTCTCTTTTGGCGGGGCACTGTGCGTGGAAGC 300  
Qy 534 GTAGACAAGGAGATGCAAGGTATTTGGTCGATTCGATTCGATTCGATTCGATTCGATTCG 593  
Db 301 GTAGACAAGGAGATGCAAGGTATTTGGTCGATTCGATTCGATTCGATTCGATTCGATTCG 360  
Qy 594 AATGACCACTTAGAGCTTTGGATCCAGAGAACCGCGCTGGGACACTTTTGTGGATCTC 653  
Db 361 AATGACCACTTAGAGCTTTGGATCCAGAGAACCGCGCTGGGACACTTTTGTGGATCTC 420  
Qy 654 TACGGGAACCAATGACAGCGCGGAGAGCGGGAAGAGCCAGGAGCGTTTCAACCGCTGGTTC 713  
Db 421 TACGGGAACCAATGACAGCGCGGAGAGCGGGAAGAGCCAGGAGCGTTTCAACCGCTGGTTC 480

ORIGIN

Query Match 42.6%; Score 742.6; DB 7; Length 767;  
Best Local Similarity 99.2%; Pred. No. 3.9e-192;  
Matches 757; Conservative 0; Mismatches 2; Gaps 1;  
Qy 234 CCATCTCTGGCACCTGGCGGATAGCCCGCGGTGAATGGAGCCACTGGCCACAGCAGCAGT 293  
Db 1 CCATCTCTGGCACCTGGCGGATAGCCCGCGGTGAATGGAGCCACTGGCCACAGCAGCAGT 60  
Qy 294 TTGATGCGGGGAGGTAAATCCCATGCGCAGCATGTAAGCAAGCGCTGAGAGAGCTGGC 353  
Db 61 TTGATGCGGGGAGGTAAATCCCATGCGCAGCATGTAAGCAAGCGCTGAGAGAGCTGGC 120  
Qy 354 GATCAGTTTGAACCTGCGGTACCGGAGAGCATTCAGTCGATCTTAACATCCAGCTTCATATA 413  
Db 121 GATGAGTTTGAACCTGCGGTACCGGAGAGCATTCAGTCGATCTTAACATCCAGCTTCATATA 180  
Qy 414 ACCCCAGGACAGCATATCAGAGCTTTTGAACAGAGTAGTGAATGAATCTTTTCGGGATGGG 473  
Db 181 ACCCCAGGACAGCATATCAGAGCTTTTGAACAGAGTAGTGAATGAATCTTTTCGGGATGGG 240  
Qy 474 GTAAACTGGGGTGCATTTGGCCCTTCTCTTTTGGCGGGGCACTGTGCGTGGAAGC 533  
Db 241 GTAAACTGGGGTGCATTTGGCCCTTCTCTTTTGGCGGGGCACTGTGCGTGGAAGC 300  
Qy 534 GTAGACAAGGAGATGCAAGGTATTTGGTCGATTCGATTCGATTCGATTCGATTCGATTCG 593  
Db 301 GTAGACAAGGAGATGCAAGGTATTTGGTCGATTCGATTCGATTCGATTCGATTCGATTCG 360  
Qy 594 AATGACCACTTAGAGCTTTGGATCCAGAGAACCGCGCTGGGACACTTTTGTGGATCTC 653  
Db 361 AATGACCACTTAGAGCTTTGGATCCAGAGAACCGCGCTGGGACACTTTTGTGGATCTC 420  
Qy 654 TACGGGAACCAATGACAGCGCGGAGAGCGGGAAGAGCCAGGAGCGTTTCAACCGCTGGTTC 713  
Db 421 TACGGGAACCAATGACAGCGCGGAGAGCGGGAAGAGCCAGGAGCGTTTCAACCGCTGGTTC 480



```

Qy 714 CTGACGGGCGATGACTGTGGCTGTGTAGTTCTGCTGGGCTCACTTTCACTGCGGAAGTGA 773
Db 481 CTGACGGGCGATGACTGTGGCTGTGTAGTTCTGCTGGGCTCACTTTCACTGCGGAAGTGA 540
Qy 774 CCAGACACTGACGGTGCACCTCTCACTCTCCACCTTGCCTCCACCACTCACTCTCTC 833
Db 541 CCAGACACTGACGGTGCACCTCTCACTCTCCACCTTGCCTCCACCACTCACTCTCTC 600
Qy 834 TTACGCCACCACTTCTTACCAAGGAGAACCACTACATGCAACTCACTGCACTCCCTTCCCTATT 893
Db 501 TTACGCCACCACTTCTTACCAAGGAGAACCACTACATGCAACTCACTGCACTCCCTTCCCTATT 660
Qy 894 TAGGTTGGGCTAGACGGAGTCCCTGCGAGTATCTTCTAGAACTTACACAGCTTCTG 953
Db 661 TAGGTTGGGCTAGACGGAGTCCCTGCGAGTATCTTCTAGAACTTACACAGCTTCTG 720
Qy 954 TGAAGCCACCTTCCCTCCACATCTCAGTTCCCTTGGCTCAA 996
Db 721 TGAAGCCACCTTCCCTCC--CATCTCAGTTCTCTTGGCTCAA 761

RESULT 3
BUI56982 1087 bp mRNA linear EST 04-SEP-2002
LOCUS AGENCOURT 7186902 NIH_MGC_100 Homo sapiens cDNA clone IMAGE:6001882
5', mRNA sequence.
ACCESSION BUI56982
VERSION BUI56982.1 GI:22670514
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1087)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: CGAP (Stanford)
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2269 row: d column: 11
High quality sequence stop: 721.
Location/Qualifiers
1..1087
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6001882"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_100"
/organism="Homo sapiens"
/site="pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Size-selected >500bp for average insert size
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."

FEATURES
source
Query Match 41.8%; Score 727.4; DB 5; Length 1087;
Best Local Similarity 82.2%; Pred. No. 6.3e-188;
Matches 901; Conservative 0; Mismatches 174; Indels 21; Gaps 5;
Qy 140 CTGGAGTCAGTTAGCGATGTGGAAGAGAACAGGACTGAAGCCCCAGAGAACTGAACC 199

```

```

Db 1 CTGGAGTCAGTTAGTGATGTGGAAGAGAACAGGACTGAAGCCCCAGAGGGACTGAATC 60
Qy 200 AGAAAGGGAGACCCCACTGAGTCCATCAATGCAACCCATCTCTGGCACCCTGGCGGATAGCCC 259
Db 61 GGAGATGGAGACCCCACTGAGTCCATCAATGCAACCCATCTCTGGCACCCTGGCAGACGCC 120
Qy 260 CGCGGTGAATGGAGCCCACTGGCCACAGCAGAGTGTGGATGCGGGGAGGTATATCCCAT 319
Db 121 CGCGGTGAATGGAGCCCACTGGCCACAGCAGAGTGTGGATGCGGGGAGGTATATCCCAT 180
Qy 320 GGCAGAGTCAAGCAAGCGCTGAGAGAGGCTGGCGATGAGTTTGAACCTGCGGTACCGGAG 379
Db 181 GGCAGAGTCAAGCAAGCGCTGAGAGAGGCTGGCGATGAGTTTGAACCTGCGGTACCGGAG 240
Qy 380 AGCATTCAGTGATCTTAACATCCAGCTTTCATATAACCCAGGGACAGCATATCAGAGCTT 439
Db 241 GGCATTCAGTGATCTTAACATCCAGCTTTCATATAACCCAGGGACAGCATATCAGAGCTT 300
Qy 440 TGAACAGGTAGTGAATGAATCTTTTCGGGATGGGGTAAACTGCGGCTGCATATGCGGCTT 499
Db 301 TGAACAGGTAGTGAATGAATCTTTTCGGGATGGGGTAAACTGCGGCTGCATATGCGGCTT 360
Qy 500 CTTCCTCTTTGGCGGGCACTGCTGGAAGGCTGAGACAGGAGATGCAAGTATTTGGT 559
Db 361 TTCTCTCTTTGGCGGGCACTGCTGGAAGGCTGAGACAGGAGATGCAAGTATTTGGT 420
Qy 560 GAGTCGGATGCAAGTTGGATGGCCACTCTCAATGAATGACCCTAGAGCTTGGATCCA 619
Db 421 GAGTCGGATGCAAGTTGGATGGCCACTCTCAATGAATGACCCTAGAGCTTGGATCCA 480
Qy 620 GGAGAACGGCGGCTGGGACACTTTTGTGGATCTCTACGGGAACAATCAGCAGCCGAGAG 679
Db 481 GGAGAACGGCGGCTGGGACACTTTTGTGGATCTCTACGGGAACAATCAGCAGCCGAGAG 540
Qy 680 CCGAAGAGGCGCAGAGCGTTTCAACCGCTGTTCTCTGACGGGCACTGCTGGCTGGTGT 739
Db 541 CCGAAGAGGCGCAGAGCGTTTCAACCGCTGTTCTCTGACGGGCACTGCTGGCTGGTGT 600
Qy 740 AGTTCTCTGGGCTCACTCTTCACTCGGAGTACACAGACTGACCTCCCTCACTCACTC 799
Db 601 GGTTCCTGCTGGGCTCACTCTTCACTCGGAGTACACAGACTGACCTCCCTCACTCACTC 660
Qy 800 TCACCTCCCACTTGGCCCCCAACCACTCTCTCTTCACTCGGACCACTGCTTACCAAGGAA 859
Db 661 TCCACACCCCTTCTCTGCTCCACCATCTCTCCCTCCAGCCGCCATTGGCCACCAAGGAA 720
Qy 860 CCACATACATGCACTCAAGCCCTTCCCTCTTATATAGGTTGGGCTTAGACGGAGTCCCC 919
Db 721 CCACATACATGCAAGCCCTTCCCTCTTATATAGGTTGGGCTTAGACGGAGTCTGCTCCT 780
Qy 920 TGCACTAGCTTTTCTAGAACTACACCGCTTCTGTGAAGCCACTTTCCTCCCACTCCTC 979
Db 781 TGCACTAGCTTTTCTAGAACTTATACACTTCTCTGTAGA-----CCCCCACTCCTC 831
Qy 980 AGTTCCCTTGGCTCAAACTCAAGGTTTCTCTCAGATCAGCTCTCTGGAGGCTGGC 1039
Db 832 AGTTCCCTTGGCTCAAGATTCACAA---ATTTCACAAATCTGTCCAAAGGAGCTGGC 888
Qy 1040 AGGAGTGGGAAGGG---GTGTGCTAGAGGGAGAGAGCTCTCTCTTGTGTGGGACCTGT 1096
Db 889 AGGTATGGGAAGGGTTGTGTGCTGGGCGAGAGGCGCTTACCTGATGGNGGCACTT 948
Qy 1097 ATTACCCCTGAGCTCTCGGGAATGCTTTCTGGAGGGAGCTGAGAGCTCTCTAAACC 1156
Db 949 TACCCCTTAGCCTCCCTGGAAATGTTTCTGCGCAGGAGCTTGAAGATTTTCAGAAC 1008
Qy 1157 TCTTCCCCCA--GAGACTAGATTGCC----TTGGTTTGTGCTGCTGCTCAGAAATG 1210
Db 1009 TCTTCCCCGAAAGGAGACTAGATGGGCGCTTTGGNTTTGAAGGTTGGGGCTCCANAAITG 1069
Qy 1211 ATCCATTTCCCACTTCT 1226

```



/clone="IMAGE:6376849"  
/tissue\_type="normal pigmented retinal epithelium"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_43"  
/note="Organ: eye; Vector: pOTB7; Site 1: XhoI; Site 2:  
EcoRI; cDNA made by oligo-dT priming. Directionally  
cloned into EcoRI/XhoI sites using the following 5'  
adaptor: GGCACGAG(G). Library constructed by Ling Hong  
in the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH\_MGC Library. |"

## ORIGIN

Query Match 41.7%; Score 726.4; DB 5; Length 878;  
Best Local Similarity 90.9%; Pred. No. 1.1e-187;  
Matches 783; Conservative 0; Mismatches 77; Indels 1; Gaps 1;

Qy 1 CACAGAGACCCAGTGTAGTGTGTTGGACAATGGACTGGTTGAGCCCATCT 60  
Db |||||  
Qy 6 CAGAGACGAGACTCAGTGTAGTGTGTTGGACAATGGACTGGTTGAGCCCATCC 65  
Db |||||  
Qy 61 CTATTATAAATCTCTCAGAGCAACCGGAGCTGGTGTGACTTCTCTCTACAGC 120  
Db |||||  
Qy 66 CTATTATAAATGTCTCAGAGCAACCGGAGCTGGTGTGACTTCTCTCTACAGC 125  
Db |||||  
Qy 121 TCTCCAGAAAGGATACAGCTGGAGTCACTTTAGCGATGTGGAAGACAAGGACTGAAG 180  
Db |||||  
Qy 126 TTTCAGAAAGGATACAGCTGGAGTCACTTTAGTGTGAGAGACAGGACTGAGG 185  
Db |||||  
Qy 181 CCCAGAGAAACTGAACAGAAAGGAGACCCCGAGTGCATCAATGGCAACCCATCCT 240  
Db |||||  
Qy 186 CCCAGAGAGGACTGAATCGGAGATGGAGACCCCGAGTGCATCAATGGCAACCCATCCT 245  
Db |||||  
Qy 241 GGCACCTGGCGATAGCCCGCGGTGATGAGCACTGGCCACAGCAGCACTTGGATG 300  
Db |||||  
Qy 246 GGCACCTGGCAGACAGCCCGCGGTGAATGAGCCACTGGCCACAGCAGCACTTGGATG 305  
Db |||||  
Qy 301 CGCGGAGGTAAATCCCATGTCAGCAGTGAAGCAAGCGCTGAGAGAGCTGGCGATGAGT 360  
Db |||||  
Qy 306 CGCGGAGGTGATCCCATGTCAGCAGTGAAGCAAGCGCTGAGGAGCAGCGGACGAGT 365  
Db |||||  
Qy 361 TTGAACCTGGCTACCGGAGACATTCAGTGTATTAACATCCAGCTTCATATAACCCGAG 420  
Db |||||  
Qy 366 TTGAACCTGGCTACCGGAGACATTCAGTGTATTAACATCCAGCTTCATATAACCCGAG 425  
Db |||||  
Qy 421 GGACAGCATATCAGAGCTTGAACAGGAGTGAATGAATCTCTTCGGAGTGGGTAACT 480  
Db |||||  
Qy 426 GGACAGCATATCAGAGCTTGAACAGGAGTGAATGAATCTCTTCGGAGTGGGTAACT 485  
Db |||||  
Qy 481 GGGGTGCGATTTGGCTCTCTCTCTTGGCGGCGACTGTGGTGGAAAGCGTAGACA 540  
Db |||||  
Qy 486 GGGGTGCGATTTGGCTCTCTCTCTTGGCGGCGCGACTGTGGTGGAAAGCGTAGACA 545  
Db |||||  
Qy 541 AGGAGATCGAGTATTTGTGTAGTGGATGCAAGTTGGATGGCCACCTACCTGATGACC 600  
Db |||||  
Qy 546 AGGAGATCGAGTATTTGTGTAGTGGATGCAAGTTGGATGGCCACCTACCTGATGACC 605  
Db |||||  
Qy 601 ACCTAGAGCTTGGATCCAGAGAAACCGCGCTGGGACACTTTTGTGATCTCTACGGGA 660  
Db |||||  
Qy 606 ACCTAGAGCTTGGATCCAGAGAAACCGCGCTGGGATCTTTTGTGAACTCTATGGGA 665  
Db |||||  
Qy 661 ACAATGACAGCCGAGAGCCGGAAGGCGCAGGAGCGTTTCAACCGCTGGTCTCTGACGG 720  
Db |||||  
Qy 666 ACAATGACAGCCGAGAGCCGGAAGGCGCAGGAAACGCTTCAACCGCTGGTCTCTGACGG 725  
Db |||||  
Qy 721 GCATGACTGTGGCTGTGTAGTCTGCTGGGCTCACTCTTCAGTTCGGAAGTACACGAC 779  
Db |||||  
Qy 726 GCATGACTGTGGCTGTGTAGTCTGCTGGGCTCACTCTTCAGTTCGGAAGTACACGAC 785  
Db |||||  
Qy 780 ACTGACCGTCCACTCACCTCTCACCTCCGACCTTGGCCCGCCACCACTCTCTCTTACG 839  
Db |||||  
Qy 786 ACTGACCATCCACTCTACCTCCGACCTTCTCTCTGCTCCCAACACATCTCTCTCGTCCAGC 845  
Db |||||

Qy 840 CACCATTGCTACCAGGAGAAC 860  
Db |||||  
Qy 846 CGCATTGCCACCAGGAGAAC 866  
Db |||||

## RESULT 6

BQ052167

LOCUS

DEFINITION

5', mRNA sequence.

ACCESSION

BQ052167

VERSION

BQ052167.1

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 997)

NIH-MGC http://mgi.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: csaps-r@mail.nih.gov

Tissue Procurement: Dr. Daniel McVicar, DBS/NCI

cDNA Library Preparation: Rubin Laboratory

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: L1CM2117 row: 1 column: 09

High quality sequence stop: 693.

Location/Qualifiers

1..997

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:5933336"

/tissue\_type="natural killer cells, cell line"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NIH\_MGC\_106"

/notes="Organ: blood; Vector: pOTB7; Site 1: XhoI; Site 2:

EcoRI; cDNA made by oligo-dT priming. Directionally cloned

into EcoRI/XhoI sites using the following 5' adaptor:

GGCACGAG(G). Library constructed by Ling Hong in the

laboratory of Gerald M. Rubin (University of California,

Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and

Superscript II RT (Life Technologies). Note: this is a

NIH\_MGC Library."

## ORIGIN

Query Match 41.5%; Score 722.4; DB 5; Length 997;  
Best Local Similarity 85.8%; Pred. No. 1.5e-186;  
Matches 841; Conservative 0; Mismatches 126; Indels 13; Gaps 3;

Qy 78 CAGAGAACCGGAGCTGGTGTGACTTCTCTCTACAGCTCTCCCAAGAGATAC 137  
Db |||||

Qy 138 AGCTGGAGTCACTTTAGCGATGTGGAAGAACAGGACTGAAGCCCAAGAACTGAA 197  
Db |||||

Qy 61 AGCTGGAGTCACTTTAGTGTGTGGAAGAACAGGACTGAAGCCCAAGAGGACTGAA 120  
Db |||||

Qy 198 CCAGAAAGGAGACCCCGAGTGCATCAATGGCAACCCATCTCTGGCACCTGGCGATGAC 257  
Db |||||

Qy 121 TCGAGATGGAGACCCCGAGTGCATCAATGGCAACCCATCTCTGGCACCTGGCAGACG 180  
Db |||||

Qy 258 CCGCGGTGATGAGCCACTGGCCAGCAGCAGTGTGGATGCGCGGAGGTATCCCC 317  
Db |||||

Qy 181 CCGCGGTGATGAGCCACTGGCCAGCAGCAGTGTGGATGCGCGGAGGTATCCCC 240  
Db |||||

Qy 318 ATGGCAGCAGTGAAGCAAGCGCTGAGAGAGGCTGGCGATGAGTTTGAACCTGCGGTACCGG 377  
Db |||||

```
Db 241 ATGCGAGCAGTAAGCAAGCGCTGAGGAGCGAGCGAGGAGTGTGAACCTCGGTACCGG 300
Qy 378 AGAGCATTCAGTGATCTAACTCCAGCTTCATATATACCCAGGACAGCATATCAGAGC 437
Db 301 CGGGCATTCAGTGACCTGACATCCAGCTCCACATCACCCAGGAGACATATCAGAGC 360
Qy 438 TTTGAAACAGGTAGTGAATGAATCTTTTCGGGATGGGTAAACTGGGGTCGCAATGTGGCC 497
Db 361 TTTGAAACAGGTAGTGAATGAATCTTTTCGGGATGGGTAAACTGGGGTCGCAATGTGGCC 420
Qy 498 TTTCTTCTCTTTGGCGGGCCACTGTGCGTGGAAAGCGTAGACAAGGAGATCGAGTATTC 557
Db 421 TTTTCTCTCTTCGGCGGGCCACTGTGCGTGGAAAGCGTAGACAAGGAGATCGAGTATTC 480
Qy 558 GTGAGTCGGATGCGAGTTGGATGCGCACCTACCTGAATGACCACTTAGAGCCCTTGGATC 617
Db 481 GTGAGTCGGATGCGAGTTGGATGCGCACCTTACCTGAATGACCACTTAGAGCCCTTGGATC 540
Qy 618 CAGGAGAACGGCGCTGGGACACTTTTCTGGATCTCTACGGGAAACAATGACGACGCGAG 677
Db 541 CAGGAGAACGGCGCTGGGATCTTTTGTGAACTCTATGGAAACAATGACGACGCGAG 600
Qy 678 AGCGGAAAGGCGAGGAGCGTTTCAACCGCTGGTTCTTACGGGCGATGACTGTGGCTGGT 737
Db 601 AGCGGAAAGGCGAGGAGCGTTTCAACCGCTGGTTCTTACGGGCGATGACTGTGGCGGC 660
Qy 738 GTAGTTCTGCTGGGCTCACTCTTCAGTCGGAGTGACAGACACTGACCGTCCACTACC 797
Db 661 GTGGTTCTGCTGGGCTCACTCTTCAGTCGGAAATGACCACTGACCACTCCACTCTAC 720
Qy 798 TCTCAGCTCCCACTTGGCCCCCACCACAACTCTCTTTCAGCCCACTTGTCTACGAGG 857
Db 721 CCTCCACCCCTCTCTGCTCCACCATCTCTCGTCCAGCGGCAATGGCACAGGAG 780
Qy 858 AACCACTACATGCACTACGCGC--CCTTCCCTCTATTATAGGTTGGGCTTAGCGGAGT 915
Db 781 AACCACTACATGCACTACGCGCCTGCCCCACTGCCCCATCAGAGGTTGGCCAGATCTGG 840
Qy 916 CCCCTGAGTTAGCTTTCTAGAACTACACGCTTCTGTGAAGCACTTCCGCCACACA 975
Db 841 CCCCTGAGCTAAATTTCTAGAAATTAATCACAATCTCTGTGAA-----GACCCACACA 891
Qy 976 TCTCAGTTCCCTTGGCTTCAAACTCACAAGTTTTCCTCAGATCAGCTCTTGGAGGC 1035
Db 892 CCTCAGTTCCCTTGGGCTTCAAAATTTCCCA--AAATTCACAAATCTGTCCAAGGAGCT 949
Qy 1036 TGGCAGGAGTGGGAAGGGGT 1055
Db 950 GGGCAGGAGTATGGAGGGT 969
```

```
RESULT 7
BX441570
LOCUS BX441570 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
DEFINITION CS0DF019YE09 5-PRIME, mRNA sequence.
ACCESSION BX441570
VERSION BX441570.2 GI:47002848
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 989)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 15, 2003 this sequence version replaced gi:30775950.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, Cp 5706 - 91057 EVRY cedex - FRANCE
Email: sequef@genoscope.cns.fr, Web : www.genoscope.cns.fr
```

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized. Library was constructed by Life Technologies, a division of Invitrogen.

This sequence belongs to sequence cluster 9697.r  
For more information about this cluster, see  
<http://www.genoscope.cns.fr/cdna?c=CS0DF019AC05QP1&c=9697.r>.

FEATURES

Location/Qualifiers  
1..989  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DF019YE09"  
/tissue\_type="FETAL BRAIN"  
/dev\_stage="fetal"  
/clone\_lib="Homo sapiens FETAL BRAIN"  
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

ORIGIN

Query Match 41.1%; Score 716.2; DB 5; Length 989;  
Best Local Similarity 93.6%; Pred. No. 7.3e-185;  
Matches 746; Conservative 1; Mismatches 50; Indels 0; Gaps 0;  
Qy 1 CACAGACGACACCCAGTGAGTGAGCAGGTGTTTGGACAATGGACTGGTGGAGCCCATCT 60  
Db 192 CAGAGACGAGACTCAGTGAGTGAGCAGGTGTTTGGACAATGGACTGGTGGAGCCCATCC 251  
Qy 61 CTATTATAAAATGCTCTCAGAGCAACCGGAGCTGGTGGTGGTCTCTCTCTACAGC 120  
Db 252 CTATTATAAAATGCTCTCAGAGCAACCGGAGCTGGTGGTGGTCTCTCTCTACAGC 311  
Qy 121 TCTCCCAAGAAAGGATCAGCTGGAGTCAGTTTACGATGTGGAAGAGAACAGGACTGAAG 180  
Db 312 TTTCACAGAAAGGATCAGCTGGAGTCAGTTTACGATGTGGAAGAGAACAGGACTGAGG 371  
Qy 181 CCCACAGAAAGAACTGAACACAGAAAGGAGACCCCCAGTCGCATCAATGGCAACCCATCT 240  
Db 372 CCCACAGAAAGGACTGAATCGAGATGAGAGACCCCCAGTCGCATCAATGGCAACCCATCT 431  
Qy 241 GGCACCTGGCGATAGAGCCCGCGGTGAATGAGACCACTGGCCACAGCAGAGTTGGATG 300  
Db 432 GGCACCTGGCAGACAGCCCCCGGTGAATGAGACCACTGGCCACAGCAGAGTTGGATG 491  
Qy 301 CGCGGAGAGTAAATCCCATGCGCAGTCAGTGAAGCAAGCGCTGAGAGAGCTGGCGATGAGT 360  
Db 492 CCCGGAGAGTGATCCCATGCGCAGTCAGTGAAGCAAGCGCTGAGAGAGCGGACGAGT 551  
Qy 361 TTGAACCTCGGTACCGGAGAGCATTCAGTGATCTTAACATCCAGCTTCATATAACCCAG 420  
Db 552 TTGAACCTCGGTACCGCGCGGCAATTCAGTGACCTGACATCCAGCTCCACATCACCCAG 611  
Qy 421 GGACAGCATATCAGAGCTTTGAACAGGTAGTGAATGAATCTTTTCGGGATCGGGTAAACT 480  
Db 612 GGACAGCATATCAGAGCTTTGAACAGGTAGTGAATGAATCTTTTCGGGATCGGGTAAACT 671  
Qy 481 GGGGTGCGATTTGGCGCTTCTTCTCTTTCGGCGGGCACTGTCGTGGAAGAGCGTAGACA 540  
Db 672 GGGGTGCGATTTGGCGCTTCTTCTCTTTCGGCGGGCACTGTCGTGGAAGAGCGTAGACA 731  
Qy 541 AGGAGATGCAAGGTATTGGTGAGTCGGATTCGAAGTTGGATGGCCACCTTACCTGAATGACC 600  
Db 732 AGGAGATGCAAGGTATTGGTGAGTCGGATTCGAAGTTGGATGGCCACCTTACCTGAATGACC 791  
Qy 601 ACCTAGAGCTTGGATCCAGAGAACGGCGCTGGGACACTTTTGTGGATCTCTACGGGA 660  
Db 792 ACCTAGAGCTTGGATCCAGAGAACGGCGCTGGGATCTTTTGTGGATCTCTATGGGA 851  
Qy 661 ACAATGACAGCCGAGAGCGCGGAAAGGCGGAGGCGTTTCAACCGCTGGTCTCTGACGG 720



TITLE  
JOURNAL  
COMMENT

National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Daniela S. Gerhard, Ph.D.  
Office of Cancer Genomics  
National Cancer Institute / NCI  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Drs. Josef Lazar & Howard Jacob, Medical  
College of Wisconsin  
cDNA Library Preparation: Express Genomics  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: L14W15368 row: o column: 02  
High quality sequence start: 17  
High quality sequence stop: 544.  
Location/Qualifiers  
1. 758  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10116"  
/clone="IMAGE:7315132"  
/sex="both"  
/tissue\_type="Brain - Pooled from several tissues from one  
or more individuals"  
/lab\_host="DH10B Tona"  
/clone\_lib="NIH\_MGC\_254"  
/note="Organ: brain/CNS; Vector: pExpress-1; Site: 1:  
EcorV; Site 2: NotI; RNA obtained from brain tissue of 8  
wk old animal. Tissues were snap-frozen and kept at -80C  
before RNA extraction and purification (Tri-reagent  
method). cDNA was primed using oligo-dT primer:  
5'-PGACTAGTTCAGTTCGAGCGGCCGCTT)25-3' and cloned into  
the EcorV/NotI sites of pExpress-1. Size-selection >1.25kb  
resulted in an average insert size of 2.18 kb. This  
primary library is not normalized (normalized library is  
NIH\_MGC\_255) and was constructed by Express Genomics  
(Frederick, MD). Note: this is a NIH\_MGC library"

ORIGIN

Query Match 40.6%; Score 707.2; DB 7; Length 758;  
Best Local Similarity 99.2%; Pred. No. 2e-182;  
Matches 720; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

Qy 534 GTAGCAAGGAGATGTCAGGATTTGGTGTGATCGGATTCGAGTTGGATGGCCACCTACCTG 593  
Db 34 GGATCAAGGAGATGTCAGGATTTGGTGTGATCGGATTCGAGTTGGATGGCCACCTACCTG 93

Qy 594 AATGACCACTAGAGCCTTGGATCCAGAGAACGGCGCTGGACACTTTTGTGATCTC 653  
Db 94 AATGACCACTAGAGCCTTGGATCCAGAGAACGGCGCTGGACACTTTTGTGATCTC 153

Qy 654 TACGGGAACAATGACGAGCGGAGAGCGGAAAGCCAGGAGCGTTTCAACCGCTGGTTC 713  
Db 154 TACGGGAACAATGACGAGCGGAGAGCGGAAAGCCAGGAGCGTTTCAACCGCTGGTTC 213

Qy 714 CTGACGGGCATGACTGTGGCTAGTTCTGTCTGGGCTCACTCTTCAGTCGGAAGTGA 773  
Db 214 CTGACGGGCATGACTGTGGCTAGTTCTGTCTGGGCTCACTCTTCAGTCGGAAGTGA 273

Qy 774 CCAGACACTGACCGTCCACTCACTCTCACTCCACCTTGGCCGCCACCACTCTCTC 833  
Db 274 CCAGACACTGACCGTCCACTCACTCTCACTCCACCTTGGCCGCCACCACTCTCTC 333

Qy 834 TTTACGCCACCAATTGTTACAGGAGAACCACTACATGCAACTCAGCCCTTCCCTATT 893  
Db 334 TTTACGCCACCAATTGTTACAGGAGAACCACTACATGCAACTCAGCCCTTCCCTATT 393

Qy 894 TAGGTTGGGCTAGACGAGTCCCTCGAGTTAGTTCTTAGAATCTACACGCTCTG 953  
Db 394 TAGGTTGGGCTAGACGAGTCCCTCGAGTTAGTTCTTAGAATCTACACGCTCTG 453

Qy 954 TGAAGCCACCTTCCCCACATCTCAGTTCCCTTGGCTCAAACCTCAAGGTTTTTC 1013  
Db 454 TGAAGCCACCTTCCCCACATCTCAGTTCCCTTGGCTCAAACCTCAAGGTTTTTC 513

Qy 1014 CTGAGTCAGCTCCTCTGGAGGCTGGCAGGAGTGGGAAGGGGTGTCTAGAGGGAGAAG 1073  
Db 514 CTGAGTCAGCTCCTCTGGAGGCTGGCAGGAGTGGGAAGGGGTGTCTAGAGGGAGAAG 573

Qy 1074 CTTGCTTGTGTGGGACCTGTATACCTGTAGCCCTCTCGGGAATGCTTTCTGGCAG 1133  
Db 574 CTTGCTTGTGTGGGACCTGTATACCTGTAGCCCTCTCGGGAATGCTTTCTGGCAG 633

Qy 1134 GGAGCTGGAGAGCTCTCTAAACCTTTCCTCCCGCAGAGACTAGATTGCTTGTGTGATGT 1193  
Db 634 GGAGCTGGAGAGCTCTCTAAACCTTTCCTCCCGCAGAGACTAGATTGCTTGTGTGATGT 693

Qy 1194 GTGTGGCTCAGAAATGATCCATTTCCCATTTCTGTCTGTCTGGGCGGCTCTCCTT 1253  
Db 694 GTGTGGCTCAGAAATGATCCATTTCCCATTTCTGTCTGTCTGGGCGGCTCTCCTT 752

Qy 1254 CCATC 1259  
Db 753 CCATC 758

RESULT 10  
BX359456  
LOCUS  
DEFINITION BX359456 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
Clone CSOD1057YN09 5-PRIME, mRNA sequence.  
ACCESSION BX359456  
VERSION BX359456.2 GI:46287636  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1134)  
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished (2001)  
On May 5, 2003 this sequence version replaced gi:30368313.  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
was normalized. Library was constructed by Life Technologies, a  
division of Invitrogen. This sequence belongs to sequence cluster  
9697.r

For more information about this cluster, see  
http://www.genoscope.cns.fr/cdna?s=CSOD1057CG05QPl&c=9697.r.

FEATURES  
Location/Qualifiers  
1. 1134  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CSOD1057YN09"  
/tissue\_type="PLACENTA COT 25-NORMALIZED"  
/clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo(dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and cloned into the Not I and EcoR V  
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 40.6%; Score 707; DB 5; Length 1134;  
Best Local Similarity 91.5%; Pred. No. 2.5e-182;  
Matches 760; Conservative 0; Mismatches 70; Indels 1; Gaps 1;

Qy 1 CACAGAGACAGCCAGTGTGAGAGAGTGTGGACAATGGAGTGTGGAGCCATCT 60  
 Db |||||  
 213 CAGAGAGAGACTCAGTGTGAGAGAGTGTGGACAATGGAGTGTGGAGCCATCT 272  
 Qy |||||  
 61 CTATTATAAAATGTCTCAGAGCAACCGGAGCTGGTGTGACTTTCTCTCTACAGC 120  
 Db |||||  
 273 CTATTATAAAATGTCTCAGAGCAACCGGAGCTGGTGTGACTTTCTCTCTACAGC 332  
 Qy |||||  
 121 TCTCCAGAAAGGATACAGCTGGAGTCACTTTAGCGATGTGAGAGCAACAGGACTGAAG 180  
 Db |||||  
 333 TTTCCAGAAAGGATACAGCTGGAGTCACTTTAGTGTGAGAGCAACAGGACTGAGG 392  
 Qy |||||  
 181 CCCAGAGAAATCTGAACAGAGAGGAGACCCCGAGTGCATCATGAGCAACCCATCT 240  
 Db |||||  
 393 CCCAGAGAGGACTGAATCGGAGATGGAGACCCCGAGTGCATCATGAGCAACCCATCT 452  
 Qy |||||  
 241 GGCACCTGGCGATAGCCCGGAGTGAATGAGACCTGGCCAGCAGCAGTGTGGATG 300  
 Db |||||  
 453 GGCACCTGGCGAGACGCCCGGAGTGAATGAGACCTGGCCAGCAGCAGTGTGGATG 512  
 Qy |||||  
 301 CGCGGAGGTAAATCCCAATGGCAGCAGTGAAGCAAGCGCTGAGAGAGCTGGCGATGAGT 360  
 Db |||||  
 513 CCCGGAGGTGATCCCATGGCAGCAGTGAAGCAAGCGCTGAGAGAGCTGGCGATGAGT 572  
 Qy |||||  
 361 TTGAACCTGGTACCGGAGAGCATTCAGTGTATTAACATCCAGCTTCATATACCCAG 420  
 Db |||||  
 573 TTGAACCTGGTACCGGAGAGCATTCAGTGTATTAACATCCAGCTTCATATACCCAG 632  
 Qy |||||  
 421 GGACAGCATATCAGAGCTTTGACAGAGTGAATGAATCTTTCGGGATGGGGTAAACT 480  
 Db |||||  
 633 GGACAGCATATCAGAGCTTTGACAGAGTGAATGAATCTTTCGGGATGGGGTAAACT 692  
 Qy |||||  
 481 GGGGTGCAATGTGGCTCTTCTCTTGGCGGGCACTGTGCGTGAAGAGCGTAGACA 540  
 Db |||||  
 693 GGGGTGCAATGTGGCTCTTCTCTTGGCGGGCACTGTGCGTGAAGAGCGTAGACA 752  
 Qy |||||  
 541 AGGAGATCAGAGTATGTGTGAGTGGATTCGAGTGGATGGATGGATGGATGGATGACC 600  
 Db |||||  
 753 AGGAGATCAGAGTATGTGTGAGTGGATTCGAGTGGATGGATGGATGGATGGATGACC 812  
 Qy |||||  
 601 ACCTAGAGCTTGGATCCAGAGAGCGGCGCTGGGACACTTTGTGATCTCTACCGGA 660  
 Db |||||  
 813 ACCTAGAGCTTGGATCCAGAGAGCGGCGCTGGGACACTTTGTGATCTCTACCGGA 872  
 Qy |||||  
 661 ACAATGACAGCCAGAGAGCGGAGAGCGGCTTCAACCGCTGTCTCTGACGG 720  
 Db |||||  
 873 ACAATGACAGCCAGAGAGCGGAGAGCGGCTTCAACCGCTGTCTCTGACGG 932  
 Qy |||||  
 721 GCATGACTGTGGCTGTGTGTCTGTGGCTCACTTTGAGTGGAGTGAAGTGAAGTGAAG 780  
 Db |||||  
 933 GCATGACTGTGGCTGTGTGTCTGTGGCTCACTTTGAGTGGAGTGAAGTGAAGTGAAG 991  
 Qy |||||  
 781 CTGACCTGCACTCACTCTCACTCCACCTTGGCCCCCAGCAGCAACTCTC 831  
 Db |||||  
 992 CTGACCTGCACTCTCACTCCACCTTCTCTGCTCCACACATCTCCGTC 1042

RESULT 11  
 CK000319 798 bp mRNA linear EST 26-NOV-2003  
 LOCUS AGNCOURT 16368639 NIH\_MGC\_220 Homo sapiens cDNA clone  
 DEFINITION IMAGE:30706817 5', mRNA sequence.  
 CK000319  
 ACCESSION CK000319.1 GI:38526353  
 VERSION EST.  
 KEYWORDS Homo sapiens (human)  
 SOURCE EST.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 798)  
 REFERENCE NIH-MGC http://mgi.nci.nih.gov/  
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
 TITLE Unpublished (1999)  
 JOURNAL

COMMENT

Contact: Daniela S. Gerhard, Ph.D.  
 Office of Cancer Genomics  
 National Cancer Institute / NIH  
 Bldg. 31 Rm10A07 Bethesda, MD 20892  
 Email: cga@nci.nih.gov  
 Tissue Procurement: James Martin, University of Iowa  
 cDNA Library Preparation: M. Bento Soares, University of Iowa  
 DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov

Plate: NDAM1070 row: g column: 18

High quality sequence stop: 703.

FEATURES

source

1..798  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:30706817"  
 /lab\_host="DH10B Tona"  
 /clone\_lib="NIH\_MGC\_220"

/note="Organ: mixed; Vector: pYX-Asc; Site 1: EcoRI;  
 Site 2: NotI; Library is oligo-dT primed and directionally  
 cloned. Denatured RNA was size fractionated on a 1% agarose  
 gel. First strand cDNA synthesis was primed with oligo-dT  
 primer containing a Not I site. Double strand cDNA was  
 size selected according to mRNA size fraction, ligated with  
 EcoR I adaptor, digested with Not I and then cloned  
 directionally into pYX-Asc vector. Average insert size  
 0.5-1kb. Adaptors 5' (AATTCGACGAGG)3' and 5'  
 (CTCTGTGGC)3'. 3' Linker sequence - GCGGCCGTGAGACC T18.  
 Sequencing primers 3' end: T3 promoter primer 5'  
 (ATTAAACCTCACTAAAGGGA)3'. 5' End: T7 promoter primer 5'  
 (TAATACGACTCACTATAGG)3'. Library was constructed in the  
 laboratory of M. Bento Soares. Average insert size 3-4kb  
 Note: this is a NIH\_MGC Library."

ORIGIN

Query Match 40.4%; Score 704.4; DB 7; Length 798;  
 Best Local Similarity 92.9%; Pred. No. 1.2e-181;  
 Matches 738; Conservative 0; Mismatches 56; Indels 0; Gaps 0;  
 Qy 1 CACAGAGACAGCCAGTGTGAGAGAGTGTGGACAATGGAGTGTGGAGCCATCT 60  
 Db |||||  
 1 CACAGAGAGACTCAGTGTGAGAGAGTGTGGACAATGGAGTGTGGAGCCATCT 60  
 Qy 61 CTATTATAAAATGTCTCAGAGCAACCGGAGCTGGTGTGACTTTCTCTCTACAGC 120  
 Db |||||  
 61 CTATTATAAAATGTCTCAGAGCAACCGGAGCTGGTGTGACTTTCTCTCTACAGC 120  
 Qy 121 TCTCCAGAAAGGATACAGCTGGAGTCACTTTAGCGATGTGAGAGCAACAGGACTGAAG 180  
 Db |||||  
 121 TTTCCAGAAAGGATACAGCTGGAGTCACTTTAGTGTGAGAGCAACAGGACTGAGG 180  
 Qy 181 CCCAGAGAAACTGAACACAGAAAGGAGAGCCCCAGTGCATCAATGGCAACCCATCT 240  
 Db |||||  
 181 CCCAGAGAGGACTGAATCGGAGATGGAGACCCCGAGTGCATCAATGGCAACCCATCT 240  
 Qy 241 GGCACCTGGCGATAGCCCGGAGTGAATGAGAGCCACTGGCCACAGCAGCAGTGTGGATG 300  
 Db |||||  
 241 GGCACCTGGCGAGACAGCCCGGAGTGAATGAGAGCCACTGGCCACAGCAGCAGTGTGGATG 300  
 Qy 301 CGCGGAGGTAAATCCCATGGCAGCAGTGAAGCAAGCGCTGAGAGAGCTGGCGATGAGT 360  
 Db |||||  
 301 CGCGGAGGTGATCCCATGGCAGCAGTGAAGCAAGCGCTGAGAGAGCTGGCGATGAGT 360  
 Qy 361 TTGAACCTGGTACCGGAGAGCATTCAGTGTATTAACATCCAGCTTCATATACCCAG 420  
 Db |||||  
 361 TTGAACCTGGTACCGGAGAGCATTCAGTGTATTAACATCCAGCTTCATATACCCAG 420  
 Qy 421 GGACAGCATATCAGAGCTTTGACAGAGTGAATGAATCTTTCGGGATGGGGTAAACT 480  
 Db |||||



Db 421 GGACAGCATATCAGAGCTTTGAAACAGGTAGTGAATGAATCTCTTCGGGATGGGGTAAACT 480  
Qy 481 GGGGTGCGATGTGGCTTCTCTCTCTTGGGGGCACTGCGTGGAAAGCGTAGACA 540  
Db 481 GGGGTGCGATGTGGCTTCTCTCTCTTGGGGGCACTGCGTGGAAAGCGTAGACA 540  
Qy 541 AGGAGATCAGAGTATTGGTGTAGTGGATTTGCAAGTTGGATGGCCACCTTACCTGAATGACC 600  
Db 541 AGGAGATCAGAGTATTGGTGTAGTGGATTCGATCGAGCTTGGATGGCCACTTACCTGAATGACC 600  
Qy 601 ACCTAGAGCTTTGGATTCAGAGAAACGGCGGTGGGAACATTTTGTGTGATCTCTACCGGGA 660  
Db 601 ACCTAGAGCTTTGGATTCAGAGAAACGGCGGTGGGAACATTTTGTGTGATCTCTATGGGA 660  
Qy 661 ACAATGACAGCGAGAGCGGAAAGCCAGGAGCGTTTCAACGCTGGTTCTCTGACGG 720  
Db 661 ACAATGACAGCGAGAGCGGAAAGCCAGGAGCGTTTCAACGCTGGTTCTCTGACGG 720  
Qy 721 GCATGACTGTGGCTGTAGTTCCTCTGGGCTCACTCTTCAGTGGGAAGTGACACAGACA 780  
Db 721 GCATGACTGTGGCGGGCTGGTTCCTCTGGGCTCACTCTTCAGTGGGAATGACACAGACA 780  
Qy 781 CTGACCGTCCACTC 794  
Db 781 CTTGACATCCACTC 794

## RESULT 12

BQ924197  
LOCUS  
DEFINITION BQ924197.1 GI:22339228 922 bp mRNA linear EST 20-AUG-2002  
5' mRNA sequence.

ACCESSION BQ924197.1 GI:22339228

VERSION  
KEYWORDS

SOURCE Homo sapiens (human)

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

## REFERENCE

NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.

Email: sgapbs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory

DNA Sequencing Arrayed by: The I.M.A.G.E. Consortium (LLNL)

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLCM2656 row: 1 column: 22

High quality sequence stop: 627.

Location/Qualifiers

1. 922

## FEATURES

## source

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6464133"  
/tissue\_type="epidermoid carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_101"  
/note="Organ: lung; Vector: pOTB7; Site: 1: EcoRI; Site 2:  
XhoI; cDNA made by oligo-dT priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GGCAGAG(G). Library constructed by Ling Hong in the  
laboratory of Gerald M. Rubin (University of California,  
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
Superscript II RT (Life Technologies). Note: this is a  
NIH\_MGC Library."

## ORIGIN

BU053776 732 bp mRNA linear EST 26-AUG-2002  
LOCUS  
DEFINITION UI-M-PD0-bzf-b-13-0-UI.r1 NIH\_BMAP\_PD0 Mus musculus cDNA clone

Query Match 40.0%; Score 696.2; DB 5; Length 922;  
Best Local Similarity 85.6%; Pred. No. 2.2e-179;  
Matches 798; Conservative 0; Mismatches 123; Indels 11; Gaps 2;  
Qy 124 CCCAAAAAGGATACAGCTGGAGTCACTTTAGTCGATGTCGAAGAACAGGACTGAAGCCC 183  
Db 1 CCCAAAAAGGATACAGCTGGAGTCACTTTAGTCGATGTCGAAGAACAGGACTGAAGCCC 60  
Qy 184 CAGNAGAACTGACACAGAAAGGAGACCCCGAGTGCATCAATGGCAACCCATCTCTGGC 243  
Db 61 CAGAAGGGATGAAATGGAGATGGAGACCCCGAGTGCATCAATGGCAACCCATCTCTGGC 120  
Qy 244 ACCTGGCGGATAGCCCCGCGGTGAATGGAGCACTGGCCACAGCAGCAGATTTGATGCGC 303  
Db 121 ACCTGGCAGACAGCCCCGCGGTGAATGGAGCCACTGGCCACAGCAGCAGATTTGATGCGC 180  
Qy 304 GGGAGGTAAATCCCATCGCAGCAGTGAAGCAAGCGCTGAGAGAGCGTGGCGATGAGTTTG 363  
Db 181 GGGAGGTGATCCCATGGCAGCAGTAAAGCAAGCGCTGAGGAGGCGCAGGAGTTTG 240  
Qy 364 AACTGCGGTACCGGAGAGCAATTCAGTGAATCAATCCAGCTTCATATACCCCGAGGA 423  
Db 241 AACTGCGGTACCGGCGGCGCATTCAGTGAATCCAGCTTCATATACCCCGAGGA 300  
Qy 424 CAGCATATCAGAGCTTTGAAACAGGTAGTGAATGAATCTTTTCGGGATGGGGTAAACTGGG 483  
Db 301 CAGCATATCAGAGCTTTGAAACAGGTAGTGAATGAATCTTTTCGGGATGGGGTAAACTGGG 360  
Qy 484 GTCCATTTGTGGCTTCTTCTCTTTGGCGGGCACTGTGCGTGGAAAGCGTAGACAAGG 543  
Db 361 GTCCATTTGTGGCTTCTTCTCTTTGGCGGGCACTGTGCGTGGAAAGCGTAGACAAGG 420  
Qy 544 AGATGACAGTATTGGTCACTCGGATTCGAAGTTGGATGGCCACCTACTGATGACCAACC 603  
Db 421 AGATGACAGTATTGGTCACTCGGATTCGAAGTTGGATGGCCACCTACTGATGACCAACC 480  
Qy 604 TAGAGCTTGGATCCAGAGAACGGCGGCTGGGACACTTTTGTGGATCTCTTACCGGAAACA 663  
Db 481 TAGAGCTTGGATCCAGAGAACGGCGGCTGGGATCTTTTGTGGATCTCTTATGGGAACA 540  
Qy 664 ATGACAGCGCGAGAGCCGGAAGGCCAGGAGCGTTTCAACCGCTGGTTCTGACGGGCA 723  
Db 541 ATGACAGCGCGAGAGCCGGAAGGCCAGGAGAACGCTTCAACCGCTGGTTCTGACGGGCA 600  
Qy 724 TGACTGTGGCTGGTGTAGTTCTGCTGGGCTCACTCTTCACTGCGAAGTGACGACACTG 783  
Db 601 TGACTGTGGGCGGCTGGTTCTGCTGNGCTCACTCTTCACTGCGAAGTGACGACACTG 660  
Qy 784 ACCGTCCACTCACCTCTCACCTCCACCTTGGCCCGCCACCACTCTCTCTTACGCCACC 843  
Db 661 ACCATCCACTTACCTTCCACCCCTTCTGCTGTACCACTCCCTC-GTCCAGCGGCC 719  
Qy 844 ATTGCTACCGAGGAACCACTTACATGCAACTCAGCGCCCTTCCCTTATATAGGGTTGGG 903  
Db 720 ATTGCCACCANGAAGCACTTACATGACGCCCATGCCCCACCTGCCCCATCAGCGGNTGG 779  
Qy 904 CCTAGACGGAGTCCCTGCGAGTGTAGCTTTCTAGAACTTACACCGCTTCTGTGAAGCCAC 963  
Db 780 CCCAGATCTGGTCCCTTGCAGNCTAGTTTCTAGAAATTTATCACACTTCTGTGAGA----- 834  
Qy 964 CTTCCCCCCACATCTCAGTTTCCCTTGGCCTCAAAACTCACAAGGTTTTTCTCTCAGATCAG 1023  
Db 835 -----CCCCACACTCAGTTCCCTTGGGCTCAGAAATTTCCACAAATATCTGT 889  
Qy 1024 CTCCTTGGAGGCTGGCAGGAGTGGGAAGGGGT 1055  
Db 890 CCAAAGGAGGCTGGCGGAGGAATGGGAAGGGTT 921

## RESULT 13

BU053776

LOCUS

DEFINITION



```

IMAGE:6402732 5', mRNA sequence.
BU053776
BU053776.1 GI:22493853
EST.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cspbs-remail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pYX-5.
Location/Qualifiers
1. 732
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6402732"
/tissue_type="whole brain"
/dev_host="embryo 12.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_FDO"
/note="Organ: brain; Vector: pYX-Asc; Site: 1: EcoR I;
Site 2: Not I; The library was constructed according to
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured mRNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with an
oligo-dT primer containing a Not I site. Double stranded
cDNA was size selected according to mRNA size fraction,
ligated with EcoR I adaptor, digested with Not I, and then
cloned directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA
tail, is TCAGAGACC. This library was created for the
University of Iowa Mouse Brain Molecular Anatomy Project
(BMAP): 'Gene Discovery in the Developing Mouse Nervous
System', supported by National Institutes of Mental Health
(NIMH), Hemin Chin, Ph.D., program coordinator."
ORIGIN
Query Match 39.8%; Score 693.6; DB 5; Length 732;
Best Local Similarity 96.7%; Pred. No. 1.1e-178;
Matches 708; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
QY 28 GTGTTTGGACAATGACTGTGAGCCCATCTCTATTATATAAAATGCTCAGAGCAACC 87
Db 1 GTGTTTGGACAATGACTGTGAGCCCATCTCTATTATATAAAATGCTCAGAGCAACC 60
QY 88 GGGAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGC 147
Db 61 GGGAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGC 120
QY 148 AGTTTAGCGATGTCGAAGAGAACAGGACTGAAGCCCGCCAGAGAACTGAACACAGAGGG 207
Db 121 AGTTTAGCGATGTCGAAGAGAACAGGACTGAAGCCCGCCAGAGAACTGAACACAGAGGG 180
QY 208 AGACCCCGAGTGCATCAATGGCAACCATCTGGCACCTGGCGATAGCCCGGCTGA 267
Db 181 AGACCTCGAGTGCATCAATGGCAACCATCTGGCACCTGGCGATAGCCCGGCTGA 240
QY 268 ATGAGGCCACTGGCCACAGCAGCAGTGTGGATGCGGGGAGGTAATCCCATGGCAGCAG 327

```

Db 241 ATGAGGCCACTGGCCACAGCAGCAGTGTGGATGCGGGAGGTGATTTCCCATGGCAGCAG 300  
 QY 328 TGAAGCAAGCGCTGAGAGAGGCTGCGGATGAGTTTGAACATGCGGTACCGAGAGCATTCA 387  
 Db 301 TGAAGCAAGCGCTGAGAGAGGCGGATGAGTTTGAACATGCGGTACCGAGAGCGTTCA 360  
 QY 388 GTGATCTAAACATCCAGCTTCATATAAACCAGGAGCAGCATATCAGAGCTTTGAACAGG 447  
 Db 361 GTGATCTAAACATCCAGCTTCATATAAACCAGGAGCAGCATATCAGAGCTTTGAACAGG 420  
 QY 448 TAGTGAATGAACCTTTTCGGGATGGGTAACCTGCGGATGGGTCGATCGTGGCCCTTTTCTCT 507  
 Db 421 TAGTGAATGAACCTTTTCGGGATGGGTAACCTGCGGATGGGTCGATCGTGGCCCTTTTCTCT 480  
 QY 508 TTGCGCGGGCAGCTGTGGTGGAAAGCGTAGACAAGGAGATCAGGTATTTGGTGGTTCGGA 567  
 Db 481 TTGCGCGGGCAGCTGTGGTGGAAAGCGTAGACAAGGAGATCAGGTATTTGGTGGTTCGGA 540  
 QY 568 TTGCAAGTTGGATGGCCACCTACCTGAATGACCACTAGAGCCCTTGGATCCAGAGAACG 627  
 Db 541 TTGCAAGTTGGATGGCCACCTATCTGAATGACCACTAGAGCCCTTGGATCCAGAGAACG 600  
 QY 628 GCGGCTGGGACACTTTTGTGATCTCTACCGGAACAATGACGAGCCGAGAGCCGGAAG 687  
 Db 601 GCGGCTGGGACACTTTTGTGATCTCTACCGGAACAATGACGAGCCGAGAGCCGGAAG 660  
 QY 688 GCCAGAGCGCTTCAACCGCTGTTCTGCGGATGAGTGTGGTGGTGGTGGTGGTGGTGGT 747  
 Db 661 GCCAGAGCGCTTCAACCGCTGTTCTGCGGATGAGTGTGGTGGTGGTGGTGGTGGTGGT 720  
 QY 748 TGGGCTCACTCT 759  
 Db 721 TGGGCTCACTCT 732

RESULT 14  
 BI409987/c  
 LOCUS  
 DEFINITION 602962839f1 NCI\_CGAP\_Lu33 Mus musculus cDNA clone IMAGE:5118513 5',  
 mRNA sequence.  
 ACCSSION BI409987  
 VERSION BI409987.1 GI:15170910  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NIH-MGC http://mgc.nci.nih.gov/.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cspbs-remail.nih.gov  
 Tissue Procurement: Gilbert Smith, Ph.D.  
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima  
 Bonaldo, Ph.D.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone Distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLAM11288 row: m column: 10  
 High quality sequence start: 7  
 High quality sequence stop: 811.  
 Location/Qualifiers  
 1. 1035  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="CZECH II"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:5118513"  
 /tissue\_type="pooled lung tumors"

FEATURES  
 source



```
|||||
Db      202 GGATTGCAAGTTGGATGGCCACCTTACCTGATGATGACCACTTAGAGCCTTGGATCCAGGAGA 261
Qy      625 ACGCGCGCTGGGACACTTTTGTGGATCTCTACGGGAACAATGACAGCCGAGAGCCGGGA 684
Db      262 ACGCGCGCTGGGACACTTTTGTGGATCTCTACGGGAACAATGACAGCCGAGAGCCGGGA 321
Qy      685 AAGCCACAGGAGCGTTTCAACCGCTGGTCTCTGACGGGATGACTGTGGCTGGTGTAGTTC 744
Db      322 AAGCCACAGGAGCGTTTCAACCGCTGGTCTCTGACGGGATGACTGTGGCTGGTGTAGTTC 381
Qy      745 TGCTGGGCTCACTCTTCAAGTCGGAAGTGACACAGACACTGACCGTCCACTCACCTCTCACCC 804
Db      382 TGCTGGGCTCACTCTTCAAGTCGGAAGTGACACAGACACTGACCGTCCACTCACCTCTCACCC 441
Qy      805 TCCCACTTGGCCCCACCACTCTCTCTTCAAGTCGGAAGTGACACAGACACTGACCGTCCACTCACCTCTCACCC 864
Db      442 TCCCACTTGGCCCCACCACTCTCTCTTCAAGTCGGAAGTGACACAGACACTGACCGTCCACTCACCTCTCACCC 501
Qy      865 ACATGCAACTCACGCCCCCTTCCCTTATATAGGGTTGGGCTTAGACGGAGTCCCTGCGAG 924
Db      502 ACATGCAACTCACGCCCCCTTCCCTTATATAGGGTTGGGCTTAGACGGAGTCCCTGCGAG 561
Qy      925 TTAGCTTTCTAGAACTACGAGCTTCTGTGAAAGCCACCTTCCCTCCCAATCTCAGTTC 984
Db      562 TTAGCTTTCTAGAACTACGAGCTTCTGTGAAAGCCACCTTCCCTCCCAATCTCAGTTC 621
Qy      985 CTTGGGCTCAAACTCACAGGTTTTTCTCAGATCAGCTCTTGGAGGCTGGCAGGAG 1044
Db      622 CTTGGGCTCAAACTCACAGGTTTTTCTCAGATCAGCTCTTGGAGGCTGGCAGGAG 681
Qy      1045 TGGGAAGGGGTGTGCTAGAGGAGAGAGCGCTGCTTGTGGTGGGACCTTGATTACCC 1103
Db      682 TGGGAAGGGGTGTGCTAGAGGAGAGAGCGCTGCTTGTGGTGGGACCTTGATTACCC 738
```

Search completed: March 31, 2005, 04:21:23  
Job time : 5252 secs

George Washington